

mhmm interaction uh yeah

oral interaction in language proficiency interviews

**pro gradu-tutkielma
FM thesis**

**Department of English
University of Helsinki
29 March 2001
Thomas Toepfer**



Tiedekunta/Osasto – Fakultet/Sektion – Faculty Faciötu pf Arts		
Tekijä – Författare – Author Toepfer, Thomas Neil		
Työn nimi – Arbetets titel – Title mhmm interaction uh yeah: oral interaction in language proficiency interviews		
Oppiaine – Läroämne – Subject English Philology		
Työn laji – Arbetets art – Level MA Thesis	Aika – Datum – Month and year 2020	Sivumäärä– Sidoantal – Number of pages 140
Tiivistelmä – Referat – Abstract <p>The description and analysis of oral language use is a daunting task. In this study, I have attempted to look not only into oral language but also to probe a bit deeper into the interaction taking place between an interviewer (S) and an interviewee (Int) in a typical oral proficiency interview (OPI). An OPI has been described many times, but apparently no description of this particular S/Int interaction has been made. Nevertheless, the practitioners of the art, or skill if you will, of OPI make a variety of claims about the relative effectiveness of various strategies in testing. It would seem that most of these are primarily based on anecdotal evidence. This study may be the beginning of a look at what interaction is occurring, particularly in the Finnish setting.</p> <p>This study presents the analysis of three actual test interviews carried out by trained interviewers working for the National Certificate (NC) of Language Proficiency (of Finland). All three interviews were at the Advanced Level in the NC scaling system.</p> <p>In my study, I concentrated on the testing research of the late 1990s with its eclectic focus. I chose those factors identified in studies as most pertinent to interaction: speech act, discourse analysis, lexicality contra grammaticality, floor, back channelling, overlapping, introspection, repairs and repetitions, accommodation, and negotiation of meaning. In results differing somewhat from certain earlier studies, I identified various forms of the multiplicity of interaction which did occur in these interviews. In fact, it was shown that these OPI proceeded to a great extent under the control, initiation, and interaction of the candidates. The interviewers managed their task well, eliciting a broad and varied language sample from each candidate.</p>		
Avainsanat – Nyckelord – Keywords National Certificates of Language Proficiency, Yleiset kielitutkinnot, EFL, ESL, interaction, oral proficiency interview, language testing, English pragmatics, interviewer, test praxis, speech act, discourse		
Säilytyspaikka – Förvaringställe – Where deposited Central Campus Library		
Muuta tietoa – Övriga uppgifter – Additional information		

TABLE OF CONTENTS

acknowledgments	i
abbreviations	ii
transcription conventions	iii
1. Introduction	1
2. Background	3
2.1 History of oral testing	3
2.2 Negotiation moves	7
2.3 Lexical density	10
2.4 Negotiating for meaning	17
2.4.1 back channeling	20
2.4.2 pauses and private speech	21
2.4.3 repairs and repetitions	22
2.5 Accommodation	23
3. Materials and methods	31
3.1 Materials and subjects	31
3.2 Methods	33
3.2.1 transcription	33
3.2.2 transcription conventions	34
3.2.3 issues considered	36
4. Results and discussion	37
4.1 Basic info	37
4.2 Starting sequences	41
4.3 Lexical questions/lexical density	44

4.4 Conversation management	48
4.4.1 back channeling	48
4.4.2 overlapping	50
4.4.3 pausing	52
4.4.4 introspection	53
4.4.5 repairs and repetitions	54
4.5 Negotiation of meaning and accommodation	58
5. Implications	63
Bibliography	65
Appendix 1	
Transcripts	72
tapescript 1	73
tapescript 2	86
tapescript 3	100
Appendix 2	
National Certificates:	
scale of oral proficiency	120
Appendix 3	
National Certificates:	
interviewer training guide	122
Appendix 4	
National Certificates:	
interviewer's assessment guide	132

ACKNOWLEDGMENTS

I wish to thank especially the director and staff of the National Certificate of Language Proficiency test [Yleiset kielitutkinnnot] for their kind permission to use videotaped interviews. Kaija Kärkkäinen, director, was instrumental in granting permission for their use. Further, Sauli Takala has been my mentor, pointing me in the direction of this study and providing a staging point for the work within the framework of the National Certificates. I would like to particularly thank Ari Huhta, whose kindness was instrumental in obtaining these tapes. In addition, the work he did on the assessment guidelines and other material, and the discussions of these we have had in joint training sessions for National Certificate interviewers have been invaluable.

I wish to thank the longsuffering patience of Prof. Matti Rissanen, who continued to believe that this work would be completed.

ABBREVIATIONS

+J	-judgmental reinforcer (backchanneling)
-J	+judgmental reinforcer (backchanneling)
c-unit	independent utterance with referential or pragmatic meaning
D	display question (accommodation)
F	fronting (accommodation)
G	grammatical items (in reference to lexical density)
I	interviewee (in the tapescripts)
Int	interviewee (in the text)
Int/S	interviewee/subject (together, emphasis on interviewee)
L	lexical items (in reference to lexical density)
NBE	National Board of Education
NC	National Certificates of Language Proficiency
O	or-question (accommodation)
OPI	oral proficiency interview
PS	private speech
S	subject (interviewer/interlocutor)
S/Int	subject/interviewee (together, emphasis on subject)
SOPI	simulated oral proficiency interview
T-unit	main clause plus subordinate clauses

TRANSCRIPTION CONVENTIONS

Primary considerations

Each line of text in the transcription corresponds to one c-unit. C-units (Brock 1986: in Foster 1998: 8) are: “independent utterances which provide referential or pragmatic meaning, i.e. utterances which are meaningful though not necessarily complete.” See also discussion in chapter 2: Background.

The transcript is a table in two columns, left: speaker identification; right: utterance, i.e. c-unit.

main symbols

S	subject (interviewer/interlocutor)
I	interviewee (elsewhere in paper, Int)
[overlapping utterance, marked at the point of overlap
]	second utterance latched to first, no overlap (if not marked, but line starts immediately to the left of the column, it means the same; in other cases: the pause is marked)
=	continuity of c-unit to another line, used either because the c-unit is too long to fit on one line of the table or because of back channeling interrupting the line, which continues despite the interruption

further symbols

[]	paralingual, primarily pausing or phatic communion (markers)
[.2]	pause: time given in tenths of a second, e.g. two-tenths of a second
[breath intake]	breath intake, primarily a pausing device
[sub]	subvocalization
/-	return to normal speech after subvocalization
[???	unclear word, jumbled word
?word	slightly unclear word, e.g. ? <u>what's</u>
[?]	rising intonation
[,]	continuing intonation
[glottal stop]	glottal stop
[chuckle]	chuckle
[laugh]	laugh
[sound of smthg]	sounds, marked, made by one S/Int
[?smthg]	comments on observations from the videotape—at right edge of column
{ word }	unusual, variant pronunciation of a word which could possibly be understood otherwise
mmmmhmmm	lengthening of sound indicated by letters
com'parable	marked, unusually stressed syllable
-	word cut off, e.g. <u>wh-</u> truncated from <u>what</u>
+	sound(s) added on afterward, e.g. <u>religion + z</u>

1. INTRODUCTION

The testing of oral language, or spoken performance, is a fascinating field, and frustrating. For those of us who are practitioners of oral testing, our theoretical knowledge and our skills in “performing” testing are offset by institutional and environmental demands. Furthermore, testers often feel that they are at the juncture of competing schools of thought regarding what it is they are testing and what is being tested. Within this framework, my study is a modest attempt to determine if the practitioners of the testing of oral performance, in this case oral proficiency interview (OPI) are really interacting with the test candidates.

Theories of language use have shifted from structuralism to communicative competence, distinguishing competence from performance and viewing competence as knowledge (Chomsky 1975). Then with this distinction, assuming for a moment that it has been accepted in the academic community, there began a debate as to whether or not competence was underlying performance in some manner. Davies (1989) claims that it is impossible to nail down the concept of communicative competence, since its shapelessness, lack of clear definitions, or more accurately perhaps, definability, render it too fluid.

All of the terms involved in describing testing, e.g. performance, competence, proficiency, usage, are in a state of more or less constant flux, although none of the practitioners of testing, particularly oral testing, would be ready to do away with any of the terms, or, more importantly, any of the distinctions which those terms allow us to make. Simply put, it is difficult to describe language with language. It is even more difficult to describe language use with language use. And it is terribly hard to describe the proficiency or performance of a test candidate on an oral proficiency interview, one of the standard forms for determining levels of oral usage. But it may be even worse

to take on the problem of looking at the interviewer's performance on such an exam.

Nevertheless, I decided to study the interaction in an oral proficiency interview, paying particular attention to the interaction of the interviewer. Further, I wanted to see how what the interviewer does affects the interviewee or interview process. I am not trying to define communicative competence, other than to refer to various aspects of it brought up in various studies. In some cases I may speculate as to how these fit a particular piece of data. Some of the main considerations in communicative competence have to do with cognitive capacity or ability, behavioral "skilledness", and sociocultural factors or skill or competence.

North (2000), in his seminal volume, describes the vast complexity of setting up scales (scales are often called bands on National Certificate testing in Finland) for evaluating language use, largely for evaluating spoken language. Most of the descriptors in North's appendices describe spoken language use. The very first scale in North's classification of scales is titled "overall interaction" (2000: 386). As a matter of fact, most of North's scales (2000: 386ff) have the term interaction or transactional, the latter implying the exchange of information or phatic communion, in their titles. In that sense, there is a very firm academic basis for looking into interaction. It is not just a momentary trend, but in North's enormous analysis has emerged as the most important factor in evaluation. Therefore, it is safe to say that it is important to look into interaction.

2. BACKGROUND

In some way or another, humans have been checking up on each other's speech since the dawn of time. It is just part of the human fabric. So it is no surprise that some of the main issues involved in oral testing were delineated decades ago. Thus, some of the history of oral testing sheds light on the problematics to be encountered. After briefly looking at the history, I have chosen to continue with a discussion of what I feel are the most salient features involved in testing oral proficiency.

2.1 History of oral testing

In setting up the problematics concerning the history of oral language testing, Spolsky (1990) points out that we are often searching from the wrong source for explanations of various phenomena. Rather than an orderly advance based on theoretical premises, the actual reasons for changes in tests and testing systems are rather often external, non-theoretical, institutional, social forces. The reasons for the development of the National Certificates (NC) are found in both these areas. The tests were a fairly orderly development out of the work domain testing system developed under the National Board of Education for several years. Further, they were to meet the needs of a nation looking to enter into the European Union.

Spolsky's discussion of oral testing is based largely on the notes and on interviews with J. O. Roach, whose report in 1945 for the University of Cambridge Local Examinations Syndicate (Roach 1945: in Spolsky 1990) was apparently one of the first discussions of reliability and validity in the testing of oral proficiency.

Roach's report (1945) starts with two questions: 1. how joint examination could coordinate the standards of various oral examiners, and 2. was there a possibility of reaching further

precision in standards? Stevens (1989: in Spolsky 1990) states that the primary purpose of the University of Cambridge Local Examinations Syndicate (UCLES) was to control the instructional process, through backwash, rather than to assess proficiency. This dichotomy is addressed below.

First, there are the problems of validity. In the traditional period of testing, it is expected that raters, examiners, arrive at a mark intuitively. The dual problem of who is setting the criteria and how they are being applied remains to this day. Roach (1945) states: "It is probably, at least to some extent, the candidates who tend to set the standard in any test which has no absolute criterion." Writing of himself in the third person, Roach further says: "He believes that the candidates taught him the standard." And the summation of his experience was: "Standards of impression marking cannot be defined beforehand merely by written instructions... ." Even when one is not interested in the outcome, as such, of an oral examination, as in this study, yet it must be admitted that the perception of rating bands by trained interviewers influences their expectations regarding the interviewees. Therefore, it can be assumed, by extrapolation, that such influence or bias should be reflected rather well in the data. This remains to be seen.

Roach (1945) mentions several possible sources of contamination threatening validity. Placing these within a bit more modern framework, we have: candidate anxiety, task type open- or closed-task), test format and environment, and candidate personality. Roach notes that males examining females, and conversely, sometimes perform differently than with members of their own sex. Not only gender, but a variety of personality factors can influence the outcome. However personality may bear on S/Int interaction, it is difficult to ascertain the extent of its influence and the way this is applied in each given situation. Therefore, that remains out of the scope of this paper, with only occasional

comments on “why” one of the dyad may have reacted to the other in a particular manner.

Candidate anxiety may be exhibited in many different ways. Experienced Ss fairly often report that they sense some form of nervousness, not only in the candidates but in themselves (NC training sessions). Yet, it must be admitted that if Int does not mention fear or excitement in the dialogue, virtually all the evidence is subjective. Therefore, it remains out of the scope of this paper.

The form of interview used on the NC exams, basically OPI, is not primarily limited in any way but topic. And the topic is selected from a list by the test candidate, announced to S when Int enters the room. But when the interview begins, since S has been instructed (see appendix 3) not to have any papers or list of prepared questions to guide the work and to focus on the elicitation of the best possible language sample, the NC OPI is, then, basically an open-task. The restriction as to topic, perhaps moving the NC a bit away from the open-task mode, is due to the necessities of test security. In other words, in order to guarantee equal opportunities for all test-takers, no candidate can utilize a “prepared” speech/dialogue. Further, using a topic protocol allows for greater uniformity testee-to-testee and furthers the objectivity of the test, thus serving validity.

Spolsky (1990: 164) in summarizing the discussion of validity lays emphasis on two issues in traditional examinations. In the first place a testing body must decide what it is that is actually being tested and how that will be carried out. In the NC, this has been determined under the statements and regulations rendered by the NBE and furthered into working principles by the National Certificates Examination Board, through its research into the salient factors. And secondly, the individual interviewer (S) has to interpret these principles and put them to work in the OPI. In order to assist S in doing this, the NC arranges

training sessions for interviewers which consist of theoretical considerations, the examination of benchmarks, and hands-on practice, all followed by immediate feedback sessions and joint feedback sessions held after actual testing dates.

Roach's (1945) second focus is reliability. Reliability is defined as "the possibility of reaching agreement among the examiners." Roach found that examiners needed to have a system of control over them in order to work as a team, otherwise the evaluations varied greatly. The use of scaling over a set of variables was viewed as an efficient way to overcome the grossest difficulties. Training of examiners, the use of chief examiners, and particularly the requirements of joint examination (multiple examiners) were the fundamental methods employed to guarantee reliability. As to conversation tests, the conclusion was that there had not been adequate study made of them. They warranted further study. Today we can claim quite the same. Oral testing requires further research.

The UCLES conversation test in the 1940s focused on comprehensibility, something we now would term a global factor. The main components of comprehensibility were fluency and correctness of speech, not pronunciation and intonation. This sounds very much like a communicative approach to oral testing, although the concept of "communicative" was not actually invented at that time. Perhaps we could say that Roach foreshadowed many of the concepts which came later. Another innovation was a test with oral questions, answered in writing, thus setting the basic format for the listening comprehension test. Roach offered sensible solutions or approaches to solutions for many of the issues that remain in language testing. In this way Roach foreshadowed the development of the language testing as a field which can be said to have begun with the Washington meeting in 1960 (Spolsky: 171). It is into this framework that the NC oral examination fits.

2.2 Negotiation moves

In order to look into what happens in the unequal S/Int dyad, I need to define some basic units of the analysis of spoken language. The term "words" is difficult enough to define but I will hold off on its consideration until I take up the topic of lexical density in section 2.3. The task of transcribing a tape or videotape is daunting, even for those who have done it for years. To put it frankly, progress in transcribing proceeds rather poorly. The more one listens to a tape the more one hears, especially as one tunes in to first one feature and then another. But in order to get a starting point for transcribing the videotapes at my disposal I had to arrive at a "best economical method" of working. In order to do so, I reviewed the literature. For the outcome of this deliberation, I refer to chapter 3. For the moment, I want to look at the basic breakdown of the oral material into manageable units.

There are two basic "units" of spoken language I want to look at. For obvious reasons, the concept of sentence simply will not suffice. Previously, in various analyses, I used Hunt's (1966) definition of a T-unit: "a main clause plus whatever subordinate clauses happen to be attached or embedded within it." This was a great step forward in the transcription of taped material. However, there is a more recent definition which I have chosen to use, the c-unit. Brock's (1986) definition terms a c-unit an independent utterance which provides referential or pragmatic meaning, utterances which are meaningful although not necessarily complete. It is rather simple to see that a T-unit would not classify an elliptical clause as a unit, while the c-unit would be sensitive enough to do so. This is a great step forward in development, as is revealed in the pausing which occurs in the material. Further, much of the phatic communion taking place remains outside the realm of T-units. Yet, it must be admitted that phatic communion is one of the most important

events in the on-line processing of speech in a dyad. Such utterances as mmm can then be identified as c-units. Clearly, they are often separate speech acts. Therefore, I have chosen to use the c-unit.

Foster (1998) set up an in-classroom exchange of information-gap tasks in order to test the hypothesis that taking part in communicative language tasks assists in the learning of L2. Foster writes (1998: 1):

"The results show no clear overall effect for task type or grouping, though there was a discernible trend for dyads doing a two-way task to produce more negotiated interaction. However, it was noticeable that many students in the small groups did not speak at all, many more in both dyads and small groups did not initiate any negotiated interaction, and very few students in either setting produced any modified utterances."

These are remarkable results, indeed. I cannot help but remark that if these comments are taken seriously, it must be seen that they contradict almost all theories concerning the use of information gap tasks in L2 instruction. However, that remains outside the scope of this paper.

Foster (1998) concludes that in spite of the main theories holding in SLA "negotiating for meaning" is not the main means language learners are inclined to use when they face gaps in understanding. This concept, the negotiation of meaning when there is a functional gap in understanding, can be studied from this data according to the scheme below.

In addition, Pica's (1989) findings are significant for this study. Pairing a NNS with a NS on an information-gap task resulted in phonological, morphological, semantic, and syntactic modifications to utterances by the NNS when requested, i.e. clarification request, by the NS. This has significance for any test dyad utilizing a NS to interview a NNS.

Foster (1998: 3) emphasizes that no study has demonstrated direct relationships between the occurrence of negotiated interaction and an increase in language proficiency. Furthermore, there is the virtual impossibility of measuring the influence of negotiation on language learning. The various factors involved may well be inseparable. This means, interpreted for this data, that an Int could well demonstrate a remarkable language proficiency with very little negotiation, and conversely, another Int could have the opposite situation. The S may not necessarily pick up on this difference and may approach various Ints in an inappropriate manner. This will be returned to in the results.

In terms of negotiation moves, Foster (1998) found that almost all moves were semantic in nature, i.e. the paraphrasing of a problematic utterance or a word substitution. The data at hand can be checked to determine if this held true in these testing dyads. Foster's findings that almost no students made attempts at negotiating for meaning presents challenges to the interviewers. Furthermore, checking of interlocutors' questions to determine if, firstly, they fit into the interviewees' discourse, and, secondly, they are answered by Ints or simply ignored as was found most often in Foster (1989). The fact that the most common method of dealing with gaps, pretending to understand and hoping that the explanation will arrive soon, may be an inadequate strategy for the S. The interviewer needs to check in order to determine if there is a roof effect on the language sample or if the interviewee is simply trying to bluff.

Swain (1996) states that Output has a three-fold function: 1. it is an opportunity for language learners to notice gaps in the knowledge of the L2 that need to be filled; 2. it enables them to test out hypotheses about the structure of the L2; and 3. it allows them to reflect consciously upon the structure of the L2. Because most tasks used in research are

communicative rather than focusing on form, they approach language from the viewpoint of meaning. Whether or not one considers this the correct approach, it does not encourage reflection on the language being used. A language interview, by its nature, is rather more a meaning-based communicative event than a form-based reflection on usage. That has to be. But this factor has implications for testing. One way of dealing with this issue, lexical density, will be discussed in the following section.

In analysis Foster (1998: 8) uses the negotiation moves of Chaudron (1988), all of these being applied here and in Foster to dyads:

1. confirmation checks: A asks B whether A's verbally expressed understanding of B's utterance is correct;
2. clarification requests: A asks B for more information about B's previous utterance;
3. comprehension checks: A asks B whether B has understood A's previous utterance.

In addition to calculating sums of negotiation moves made over various test tasks, Foster (1998: 12) calculated the number of negotiated input moves as a percentage of total c-units. As these calculations yield relative sums, they can then be used for comparisons with the data at hand.

The understanding of the S/Int negotiation moves thus forms the basis of the understanding of the interaction of the interviewer and interviewee. In order to look at these moves the transcription has been made according to the appropriate c-units, with the apparatus described in chapter 3.

2.3 Lexical density

Lexical density, originally coined by Ure (1971), is a measure of the relationship between the number of words with lexical as opposed to grammatical properties as a percentage of the total words in a text. In the opening paragraph of the previous section, I stated that I would

return to the concept of “word” at this point. The core of the problem is that there are many utterances in phatic speech which are not words in the traditional sense at all. The issue is what to do with them. O’Loughlin (1995) deals with these same problems in a manner worth emulating.

O’Loughlin (1995) considers the problem of lexical density, which is founded on the count of the relative density of content to non-content words, in a study four tasks carried out in the classroom by NNSs. O’Loughlin’s (1995) analysis was carried out in conjunction with the *access:* test (the Australian Assessment of Communicative English Skills).

O’Loughlin (1995) found that the median scores across tasks fell between 35.0 and 43.0%, with a significant effect on 1. lexical density for test format and 2. lexical density for task type, but no significant effect on lexical density for interaction between test format and task type (until analysis method no. 2 was applied—see discussion below). The tape-based version produced higher lexical density than the live version. Further, the lexical density was lower on the narration task than the description and discussion tasks on both versions. In the live version, the role play had clearly lower scores than the other three tasks. These are the most important findings.

I have disregarded O’Loughlin’s second method for calculating lexical density, i.e. where the high- and low-frequency lexical items are counted differentially, the high-frequency lexical items receiving half the weight of the low-frequency items. O’Loughlin suggests this may be important for making statistical analysis, based on his own findings and the theory of Halliday (1985), but there is no reason for me to do the same, since I am not undertaking a factorial procedure on the data. Indeed, it may be safe to state that without formal reasons for undertaking such a division, it may simply be thought to be an artificial separation.

Besides being tested through various forms of tape-mediated tasks, oral usage/proficiency is most often tested by some form of OPI or SOPI, delineated as follows:

1. The OPI is a face-to-face interview by a trained interlocutor (who usually also carries out the assessment) and can include a role-play segment. The choice of questions is free, while topics and language input by S are adjusted for each Int.
2. The SOPI is a taped-medium test, invariant in nature. The SOPI is assessed afterwards by trained raters.

Stansfield (1991) proffers that one of the important difficulties in the OPI is that Int's performance is largely determined by the skill of S. The argument continues that OPI and SOPI correlate so very highly because neither allows Int to demonstrate interactive skills. This is a contention that can be demonstrated directly from the tapescripted data, i.e. whether or not this happens with any degree of regularity. Stansfield claims, quite reasonably, that in an OPI both the interviewer and the candidate understand that it is the responsibility of the candidate to perform and so very little interaction takes place (Stansfield 1991: 205). As a matter of fact, people are often paying for rather precisely this privilege. They want to show what they know. I would contend that Stansfield's claim is rather a given of the situation. However, to say that it is realized in precisely these terms may be a bit over-reaching. If back channeling, negotiation of meaning, etc. can be discovered, then it would seem reasonable to say that Stansfield's position may need a slight modification. O'Loughlin continues from Stansfield, pointing out that it is the nature of the interaction which needs to be understood. That allows us to understand what type of language is under measurement. An oral interview is a dynamic affair, with few exceptions.

Shohamy (1994) found, in a discourse analytic study of language output in OPI and SOPI, that SOPI elicited a rather more limited range of language functions; that SOPI answers

included more self-correction, e.g. face-saving, repetition of phrases in the eliciting questions and paraphrasing; that there was a restricted range of prosodic features, mainly hesitations and silence when no answer was available; and that the discourse produced in SOPI was, despite the previous results, more formal and cohesive. In addition, the grammar to lexicon ratio was 40%:60% for OPI, and just the converse for SOPI, 60%:40%. The SOPI produced language which was quite more literate. The lack of an interlocutor may have created this artifact. I found no comments in the literature that a testee may have felt that a tape produced interaction.

The term "lexical density," as it was originally coined by Ure (1971), means the measurement of the relationship between the number of words with lexical as opposed to grammatical properties as a percentage of the total words in a text. Ure found, over a wide range of written and spoken texts, that spoken texts had a lexical density usually not exceeding 40% (range 24-43%, while written texts exceeded 40% (range 36-57%). Spoken texts with feedback to the speaker had a lexical density of greater than 37%, those with feedback (e.g. interviews) less than 36%. When planning was allowed, it increased the lexical density to 37% or higher.

Halliday (1985) distinguishes between grammatical items, function words operating in closed, finite systems in the language and lexical items, content words entering into open sets which are infinitely extendable. Grammatical items include determiners, pronouns, most prepositions, conjunctions, forms of the verbs "to be" and "to have," some classes of adverbs (esp. interrogative and negative adverbs), and all proforms (not just pronouns). Lexical items include nouns and finite verbs. This leaves as problematic the classification of discourse markers (words and expressions used to structure discourse, including linking and sequencing devices), interjections, reactive tokens (well, yes, no), and

lexical and nonlexical filled pauses. For practical purposes, and following hints left in the published tapescripts in journal articles (although none of the research I reviewed stated this, yet it could be discerned from some transcripts), I am treating these as grammatical items.

Hasan (1988) found that the mean lexical density per T-unit of student NNSs was fairly high in formal (38.96%) and informal (40.64%) classroom interactions, informal classroom discussion (43.69%), and informal conversation with NSs outside the classroom (41.60%). When student NNSs were compared to EFL teachers in similar situations, the results were rather equal, i.e. formal (38.25%) and informal (41.15%) classroom interactions, informal classroom discussion (37.97%). Student NNSs had rather the same lexical density for informal conversations with NSs outside the classroom (41.60%) as did EFL teachers (42.48%). Still, all these figures pale beside the findings regarding NSs acting as interviewers (47.02%) and student NNSs as interviewees in formal interviews (33.67%). This sheds light on one specific language use difference between S and Int. Because these figures are so interesting, I will attempt to see what ratios are in the data of my study.

Two studies, Stubbs (1986) and Zora and Johns-Lewis (1989), reported higher lexical densities than did Hasan's impressive study. Zora and Johns-Lewis propose eight possible sources of variation to explain the differences in these findings: 1. a different basis for calculating lexical density, assigning items to classes; 2. the expected interruption and length of speaking turn, apparently monology over against interactiveness, can create a bias in that the speakers pitch their speech at a higher degree of lexical density; 3. the function of various component units of text; 4. self-consciousness and its relationship to self-monitoring; 5. personal attributes; 6. group attributes; 7. planning time; and 8. topic used in each instance. The first of these is, of course, the most obvious. Different

transcription systems were used in the several studies reviewed, and various methods of classifying words/utterances were employed. However, the other sources listed are by no means to be overlooked.

O'Loughlin (1995) points out the discrepancy between linguistic items and words in English. More than one word may make up an item, e.g. multiword verbs catch up with, phrasal verbs, and idioms feeling the heat, kick the ball around. It is not clear whether euphemisms, e.g. passed on to his reward, would belong to the category of idiom. On the other hand a word may have more than one item, e.g. contractions. Due to the possibilities of various meanings in a word, different items may be achieved by the same word. In addition, different words can sometimes realize one lexical item. Further, a linguistic item may be a particle, e.g. mhmm or uh.

Zora and Johns-Lewis (1989) note that the analysis of the framework should be carried out manually rather than automatically (software) in order to gain in accuracy. This allows the researcher to evaluate each item within its real context. No computer program has yet been devised which would surpass manual counting, unfortunately.

O'Loughlin (1995) counted only audible items, which may have slightly raised the lexical density reported. Most of the inaudible items appear, within context, to be mispronounced lexical items. In addition, O'Loughlin did not take into consideration all nonlexical filled pauses except where they had "a clear discourse marking function" (1995: 229). And for self-repair only the final version of an item was used in the analysis. However, I would contend that many of the inaudible, or almost inaudible, items were various types of particles. To contend that these have relatively little or no discourse function may be rather premature. As long as they actually exist in these discourse samples, we need to display them for analysis. The lack of a discourse function simply

demonstrates our current knowledge of discourse. Therefore, I have included as much of the audible, even the barely audible, material from the tapes as I possibly could.

In the live version (O'Loughlin 1995), on everything other than the live role play, i.e. the description, narration, and discussion tasks, interlocutor feedback (in this study Int feedback, interlocutor clearly has no straightforward frame of reference in the literature as of the present moment) was almost entirely limited to reactive tokens, particles. The role play resulted in clearly lower scores than the other tasks. This would indicate that the role of S on any task would influence the degree of lexical density to a considerable extent, whenever there is more interaction there is lower density. On the other hand, "open" tasks elicit language having a higher degree of lexical density than do "closed." O'Loughlin suggests (1995: 235) that Int's perception of probable response time is reflected in lexical density. Tape-based samples have greater density because the perception of that time is limited.

In addition to interactiveness, there may well be further factors involved in determining lexical density. First, interviewees may focus on the assessment of the test, whether that is done continuously during the test or following, via tape. This perception may be correct or incorrect. But it affects the output. Secondly, a focus on content may lead to a greater density. O'Loughlin sees content as affecting primarily the tape-mediated the version. I would argue that it could change the outcome of a live interview as well. If S concentrates on content-based questions at any point in the interview, I may respond. Evidence for this may, perhaps, be discovered in the data of this study. Summarizing the last two paragraphs, lexical density varies over 1. interaction, 2. Int/S-focus on interview process, and 3. Int/S-focus on content.

It may be that lexical density reveals something about the nature of the interview interaction, particularly if differences between S and Int lexical density can be found, as above.

2.4 Negotiating for meaning

Negotiating for meaning is the "process whereby second language learners confer with each other in order to achieve understanding" (Schwartz 1980: 139). This definition is actually a bit narrow. In reality negotiating for meaning is a major feature of much of NNS communication and NNS/NS or NS/NNS dialogue as well. This concept of negotiation is developed further by Brown and Levinson (1987: 13f.) as positive face (a desire of approval) and negative face (a desire to be unimpeded in one's actions).

White (1997) studied non-native participants as users rather than as learners of English. The study used a corpus of simulated sales negotiations involving American and Japanese participants, focusing on differences in 1. back channeling, 2. repair, 3. repetition, 4. pausing, and 5. private speech found in these two groups of subjects. The transcriptions were analyzed with the Longman Miniconcordancer which analyzes a wide range of features, including pausing and back channeling cues. The pauses were manually timed with a digital stop watch in hundredths of a second. These pauses then formed an important part of the study.

White (1997: 314) found:

"...the function of back channelling differs, the Americans favouring [+judgmental] reinforcers, the Japanese [-judgmental] prompters. Pausing is ... a negotiating management tactic more effectively deployed by the American participants, while repair, a feature of the Americans' behaviour, appears to have a tactical function as a marker of tentativeness in the management of face wants

and the pursuit of solidarity. It is suggested that differences in the deployment of such features by American and Japanese subjects can result in pragma-linguistic breakdown, which in turn is linked to culturally related norm differences."

White's (1997: 322-23) main forms of back channeling were such short utterances as yes, mmm, then sentence completions, and requests for clarification. The matter of the taxonomy of interactive questions will be taken up in 2.5.

Dow (1989: 45) stresses that teaching materials, negotiation simulations, need to be revised to focus "on the peculiarities of English politeness phenomena such as hedges as downgraders and features of mitigation." Teachers do at times attempt to teach this, but with varied success.

Ulijn and Gorter (1989: 5) found that "speaking a foreign language fluently is certainly a great advantage but when the foreign language speaker makes an error, the native language speaker may take it literally and, in consequence, there may be a misunderstanding." Errors of this type by S need to be considered in the study at hand.

Halmari (1993) amply demonstrated that, in business telephone conversation routines, greater prominence was given to the non-topical element of the conversation by Finnish speakers than American English speakers. The Americans interrupted more, and such overlapping tended to be in the middle of the interlocutor's turn. However, the Finns overlapped, when they did, the last element. The transcription apparatus allows for the comparison of these findings to the current study.

Graham (1985) found that Brazilians had more overlaps than Americans or Japanese. Note that in White's study (1997), the Americans had marginally fewer conversational overlaps than the Japanese (10.3 per 10 minutes for the

Americans, 12.6 for the Japanese). These statistics lend themselves to comparison with the current data.

Yamada (1992) showed that the conversational turns demonstrated important differences in regard to turn-distribution patterns. The Japanese speaking Japanese together took shorter turns which were distributed fairly evenly among the participants regardless of who initiated each topic. Among Americans in a similar situation, the person initiating a topic usually had the highest number of turns, the turns were distributed unevenly, and there were several observances of long monologic turns. These results, as such, are not directly comparable to the present study. However, the distribution of turns in an unequal dyad, tester/interviewer vs. testee, may show some trends. Further, the possibility of S (interviewer) dominating a particular segment of an interview can be revealed by the data.

There has been much discussion of the irreality of the oral test situation, discussion in the past (Roach 1945) and continuing at present. It is not likely to abate any time soon. The stakeholders in an oral interview (also in addition to the two persons in the room at the time of the interview) hold rather unequal shares. This brings up the question of the authenticity of the situation, the reality of the language sample, and the problem of ability vs. performance. However, a partial answer can be taken from March (1989), who reported that there were no substantial differences in the negotiating behavior of American and Japanese subjects in simulated and authentic negotiating events. Apparently, no one has proposed that the study could be contaminated by observer (taping) presence in authentic negotiations. After all, such negotiation/interview events may well be under observation to ensure the satisfaction of various legal requirements. For the purposes of my study, it can be inferred from the above that an oral interview has a certain authenticity with regard to the speech act and the language

sample obtained. McCreary (1986) has findings supporting those of March (1989). Furthermore, LoCastro (1990) discovered in recording an authentic meeting that the subjects did not considerably change their communication style or their language in "real" situations contra so-called artificial situations.

2.4.1. back channeling:

Hayashi and Hayashi (1991: 121ff) take the position in their definition of turn and back channel behaviors that BC is not be analyzed simply on a mechanical level but at a higher domain of cognition, entailing that BC is within the context of floor, or even topic, ownership and community competence, which they see as that social competence that the interactants need in order to participate in a contiguous conversation event, anticipating and evaluating utterances and phenomena. They propose the following taxonomy of BC:

continuers	do not require a response
repairers	call for a response (also self-repair)
reinforcers	+judgmental
claimers	+judgmental
prompters	-judgmental (non-judgmental)
clarifiers	-judgmental

Hayashi et al. (1991) state that the timing of BC, both in English and Japanese towards the end of an utterance, is important in its appropriate use. That would imply that inappropriate use would have differential timing, i.e. initial or mid-utterance BC.

Yet, one is left with the question whether the conversation in a dyad is the joint creation of the partners, Japanese, or the demonstration of antagonistic behavior, American. It must be admitted that Hayashi et al. (1991) have rather different findings. According to them, Japanese and American discourse styles do not exhibit very great differences. The

problematics of joint creation may well remain an open question for some time to come.

2.4.2. pauses and private speech

Goldman (1994: 253, 274-5) considers the contrast between the American aversion to silences of greater than 30 seconds and the Japanese use of "pregnant" silence exceeding the same time span.

Pausology focuses on speech and articulation rate and silent pause phenomena. Jefferson (1989) found over a wide range of spontaneous conversations that there was apparently a cluster of pauses of approximately 1 sec., this being the "standard maximum silence" of the "standard tolerance." After 1 sec. the pause may be interrupted. Filled pauses are realized in a variety of fashions: ah, ahm, er, etc. Among the other pause phenomena suggested are, most importantly, breath intake, then, coughing, clearing the throat, even sucking the teeth or lip.

White (1997: 330) found 18 pauses exceeding the 1 sec. standard. Of course, the fact that his Ss were negotiating over prices, calculating both in their heads and on calculators, creates a large difference. Therefore, this figure is not comparable to my study.

Private speech (PS), egocentric speech having primarily a self-regulating function, e.g. why did I say that, is defined by Lantolf and Appel (1994:15) as representing, "an externalization of the inner order as the individual attempts to regain control of his or her cognitive functioning to carry out the task." McCafferty (1994) classifies self-regulatory PS as meeting the requirements of: 1. an effort to seek self-guidance, basically and 2. mastery of some task-relevant difficulty. Furthermore it is suggested that PS may have an other-regulatory function as well. This is a factor that can be searched from my data.

2.4.3. repairs and repetitions

Repairs and repetitions are, according to White (1997), the outcome of locally managed discourse. This discourse has as its function the promotion of a successful outcome in an event that is characterized by a quite high degree of unpredictability, due to cultural differences between the participants, e.g. of a language interview dyad. Stubbs (1983) finds a great deal of phatic communion in speech, much of which is seen as being taken up by structuring, repeating, probing, emphasizing, mitigating.

Self-repair, as claimed by Schegloff, Jefferson, and Sacks (1977), is the preferred manner of repair, primarily due to its face-saving value (cf. Brown and Levinson 1987). White (1997) differentiates between repair, replacement with meaning implications, and repetition, reiteration of the same lexical item. However, this is looking at discourse from only one main viewpoint, that of propositional content. White argues that much of on-line planning during impromptu speech gives evidence of social content of the discourse, tentativeness, mitigation, self-disclosure, unfilled pauses as discourse markers for leaving the floor open.

In his conclusions, White (1997: 339) suggests that the use of discoursal features, particularly interaction schemes and BC cues, may lead to the mutual (mis)interpretation of such features according to variant cultural behavioral models. This may lead to pragmalinguistic failure or "at least, pragmalinguistic breakdown." In White's study the American interactor's attempts at mitigation may not actually lead to the avoidance or redress of imposition but may further contribute to the divergence of the dyadic discourse. This would seem to imply that in a language-test dyad, increased attempts by S to accommodate the speech of Int may lead to less, not more, understanding. This would, apparently, except it be redressed through later on-line negotiations,

lessen the quality of the speech sample obtained. However, this is simply jumping into the next section.

Stubbs (1983: 178) terms much of interaction "padding." This is not meant pejoratively but as a term of familiarity. We do pad our speech; in the on-line processing and analysis of speech we are faced with a plethora of meanings and intentions. In summary, negotiating for meaning surely reveals interaction in the S/Int dyad, particularly as regards back channeling, pauses and private speech, and repairs and repetitions.

2.5 Accommodation

Ross (1992) provides a framework for studying accommodation in oral proficiency interviews, using variable rule analyses of 16 full-length OPI. The interviews for the National Certification in Finland are in many respects, just as in all modern language interviews, fairly similar. Therefore, the data in my study lends itself to comparison. Ross (1992: 173) points out that the "gap between the widely endorsed 'proficiency movement' and a clear understanding of how oral proficiency can be assessed remains to be bridged."

As noted in the studies of Lantolf and Frawley (1985; 1988), in oral interviews the rating criteria used to define rating levels are not grounded in a set of factors which come from natural communication. This leads to the situation where it is possible to state that an interview assesses test-taking behavior as much as it does oral proficiency. Of course, researchers disagree as to which side of this equation has the larger balance. But, it must be admitted, that test-taking behavior is one of the main factors to be taken into consideration. Bachman (1988) proposes that the interview process itself results in a view of language use lacking the broader context of transactional or interactional discourse. Ross (1992) points out that our definitions of proficiency are

often derived from rather abstract and idealized views of NS-NNS interaction. To a certain extent this is confirmed by Barnwell (1989), with results showing that the ACTFL proficiency scales were not transparent enough for consistent use by nonspecialist native raters (they were not meant to be!).

In order to overcome the difficulty outlined above Ross (1992) designed a scale to measure the extent of interviewer accommodation in relation to the interviewee. I wish to do a comparison to Ross's findings in this study. This is made possible particularly through the material in the appendix to the article (Ross 1992: 185-6).

Ross (1992: 177) gives the following description of accommodative questions by interviewers:

- “1. Display question—the interviewer asks for information which is already known to the interview, or which the interviewer believes the interviewee ought to know.
2. Or-question—the interviewer asks a question and immediately provides one or more options from which the interviewee may choose an answer.
3. Fronting—the interviewer provides one or more utterances to foreground a topic and set the stage for the interviewee's response.
4. Grammatical simplification—the interviewer modifies the syntactic or semantic structure of an utterance so as to facilitate comprehension.
5. Slowdown—the interviewer reduces the speed of an utterance.
6. Over-articulation—the interviewer exaggerates the stress and production of words and phrases.
7. Lexical simplification—the interviewer chooses what is assumed to be a simpler form of a word or phrase which the interviewer believes the interviewee is unable to comprehend.”

Simply from a theoretical viewpoint, for further research, I would like to propose the following addition to Ross's scheme, item 6a, making this into an 8 point taxonomy.

6a. Flattened intonation—the interviewer flattens the intonation in order to simplify the phonemic structure, upon noticing that the interviewee is misled by various intonational patterns. This is one of those features reported in feedback sessions in the National Certificates testing interviewer training program.

Ross and Berwick (1990) found that NSs accommodate in "foreigner talk" to NNSs who seem to be incapable of carrying out their roles as conversational partners.

Interviewers who carried more than their share of the weight during interviews, that is interviewee behavior set off these 'foreigner talk' routines in trained interviewers, had accommodative, i.e. improved, impressions of the proficiency of their partners. This is discussed in the following paragraph. In addition to these results, Ross and Berwick (1990) found a greater fit between the processes of the interview and the "non-test real world."

When one applies the above to the interview process, the question comes up as to what the interviewer does upon perception of difficulties in the interview process, what signals or types of signals the interviewer receives, or what signals the interviewer gives as to accommodation. Since the interviewer is a well-trained professional, the accommodation surely takes place rather automatically. If, speaking entirely theoretically, S had to focus in each situation on the type and amount of accommodation which should be used in a particular situation, if S were forced in some way to concentrate and think out how this accommodation to the given situation should be applied, i.e. the process which S should use at this time, then the interview would be full of major pauses. The interviewer would need time, perhaps minutes, to consider each conversational move. This would relegate the interaction to

the realm of the ridiculous. Instead, an interview is a very smooth affair from the point of S. It is definitely on-line. That it is taxing of interviewers has been reported to me countless times in training sessions. I know it myself from conducting, in emergencies, up to 8 hours of interviews in a single day.

The interviewer often has years of intercultural or cultural immersion experience. In addition, the interviewer may often have extensive knowledge of and proficiency of the L1 of the interviewee. This certainly has bearing on the "foreigner language" employed by S. The on-line "foreigner language" used is, discounting exceptional situations, not a conscious effort. Therefore, its study through interviewer reports, etc. is exceptionally difficult. However, it can be studied from adequate tapescripting of actual language interviews.

Ross's (1992: 177) list of potential triggers setting off question accommodation include, in the terminology of my study:

1. Int's last response to the previous S turn,
2. foregrounding of the current discourse topic,
3. if Int made a comprehensible answer or statement during the last turn,
4. the last speaker in the previous turn, and
5. if the most recent S question had been accommodated

Ross also used the final rating given to the interviewee as factor 6, but that is not applicable to this study as the final rating was actually awarded by a reviewer working from the videotape.

Upon the application of binomial variable rule analysis of 598 interview questions, Ross (1992: 178) found that the most salient factors were, once again, in the term of my study:

1. Int's response to the previous S question,
2. the structure of Int's response to the previous S question,

3. the perceived (by S) level of Int, and
 4. if accommodation had been used in the previous question.
- Looking more closely into each of these, the first is almost entirely based on accommodated repetition by Int to a request for clarification.

First, for the purposes of this study, I would like to turn this matter on its head, so to speak. I can then address the issues of whether or not S requests for clarification are forthcoming, and how they are answered. Whenever Int gives a quick response in an appropriate on-target manner, one would expect that S accommodation would not be forthcoming.

Secondly, whenever there is a response which is phonemically variant or readily miscomprehended there is an increase in the probability of S accommodation. Ross, however, does not find any systematic use of accommodation after well-formed responses. That should perhaps be some small warning signal for interviewers. We may be over-accommodating in the sense that we accommodate when there is no apparent need. On the other hand, there is the possibility that the interviewer in the on-line situation is sensitive to something which is not available from video- or audiotape.

Thirdly, the S perceived outcome of the interview may create skew in the interpretation of need of accommodation. If S thinks from either the opening gambit or from the onset of the test that Ints will perform at a certain level, then S may well pitch accommodation to that perception. S may derive this anticipation of accommodation from, e.g., involvement within a certain research or study program or, alternatively, considerations of accent or pronunciation varieties. In a way, in the entire EFL/ESL paradigm this give and take is, naturally, what is being done—constantly. I contend that this is natural, because it is not possible in such social situations to create an outside, completely objective

reality. We interviewers attempt to remain objective and unbiased. But within the broader framework, it must be admitted there are certain scalar boundary conditions involved. Returning to the concept of S's anticipation of needed accommodation, this is a non-discourse variable, but in one way it may be regarded as a discourse variable. If one were to regard a language test itself as process, presumably meaning from the time someone came up with the idea that another language test is needed until the time that test is discarded and forgotten, then the test would become a social factor and thereby, at least marginally, a discourse factor.

Fourthly, Ross (1992: 181) finds a probability of .58 for carry-over for previously accommodated questions. This appears to be proactive accommodation, whereas that mentioned above was reactive. Short-term memory constraints afford a simple explanation of this. S must hold in mind the most important features of the interview. After a certain point in the turn-taking, it is no longer mentally economical to change the line of accommodation. In a way, this substantiates the report of many interviewers in training sessions that first impressions are vital for the interview, both within language testing and within, e.g., the job market. Nevertheless, this issue is a further factor that needs to be taken into consideration within interviews.

Ross (1992) suggests that high levels of accommodation show systematic difficulties in the interview, while low levels reveal few problems. However, if one were to consider this from the viewpoint of social interaction of the type termed the "cocktail party," requiring very much accommodation on a sophisticated level, then perhaps a further codicil should be added. There well may be various forms of accommodation which are not covered by Ross's formulation. In fact Ross does not provide a working definition of accommodation, except for oblique reference to speech accommodation theory. In Finland where the training in foreign languages is on a very high level, there

are many extremely sophisticated speakers of English. And in national examinations, it is necessary to set criteria which differentiate among these upper scale individuals. Thus, various levels of accommodation need to be addressed. However, it must be admitted that Ross is probably correct in asserting that S accommodation could well be a useful tool in writing appropriate bands for the lower end of the oral interview scale.

Ross (1992) assumes that if there is very little evidence of accommodation, then the rating of an individual performance can proceed by using bands emphasizing issues of register and style. However, the lack of accommodation may signal something else, as demonstrated by the appendix (1992: 185-6). Most S questions did not result in appropriate responses or requests for clarification. A large majority resulted in inappropriate responses. In situations where S requested clarification, Int ignored the request or gave a *non sequitur* more often than made a modified response (1992: 179). Naturally, in an ongoing interview S could not follow up on each occurrence of these inappropriate responses.

It is safe to state that accommodation remains a very important question, but much more qualitative work on actual responses and accommodation taxonomy needs to be done before quantitative analysis is applied. The problem of "interview talk" is a constant battle among among interviewers, just as "classroom talk" is for ESL/EFL practitioners. However, pointing out of specific instances of interview talk can help interviewers to recognize it and avoid it.

In this chapter, after discussing how the history of oral proficiency testing has shaped our concepts of testing and test usage, I proceeded to look into some of the main factors involved. The concentration was on the area of interaction. In addition to looking at surface strategies, I attempted to draw attention to those factors which remain behind the

scenes, as it were, discourse strategies, accommodation, and negotiating for meaning. The surface phenomena themselves are fascinating, ranging from the mhmm of phatic communion to complex patterns of interaction. Actually, entire interviews can be viewed as texts not only within a genre, oral speech, but as texts within the sociocultural framework. However, in this study I decided to concentrate particularly on interaction.

3. MATERIALS AND METHODS

3.1 Materials and subjects

This study is based on the transcription of videotapes which were actual interviews in the National Certificate testing. The director of the National Certificate, Dr Kaija Kärkkäinen, graciously lent me these videotapes for a three week period in April-early May 1998. The interviewees and interviewers had given written permission to use some these interviews, see below, as research material. Nevertheless, strict confidentiality had to be maintained. That is why the interviewees, all Finnish speaking Finns, are identified in this study according to the Finnish scheme, i.e. Matti Meikäläinen.

These videotapes are kept in a locked room at the University of Jyväskylä.

I was provided with five videotapes for use in this study, at my request randomly selected from the tapes for which candidates had given permission for their use. The videotapes contained the National Certificate interviews of thirteen candidates. The results of the National Certificate interview were not disclosed. Of these thirteen candidates, both parties of the dyad had given permission to use the material in nine cases. There were technical problems in the recordings of a further five cases, primarily difficulties with the volume of the voice track. This rendered these cases unusable. One test had, apparently, resulted in the recording of only the last couple of minutes of the interview, perhaps a mechanical failure. I was not informed of the reasons. This left three interviews available for use.

Three interviews may not sound like much material to someone who has not worked with the purpose-driven fine transcription of oral, taped or videotaped, materials. However, there are several studies in the literature which

are based on even smaller samples (cf. Pica 1987; Sengupta 1998). The transcripts contain some 5544 items, some items being greater than a word (see discussion of item in Background). The shortest of the transcriptions took over forty hours to transcribe to this level of accuracy. Yet, as I look back over the transcriptions I see slight inconsistencies. Unfortunately, it is (and was) impossible to correct these as the material is not in my possession.

These three interviews all consist of a NS/NNS dyad. And in each case NS is male, British, each different. Further, each NNS is female, Finnish, Finnish speaking (judging from name and variety of speech). Thus, there should be no gender difference per se between the three interviews (cf. Roach 1945: in Spolsky 1990). The materials, therefore, allow for no real comment on possible gender differences in S/Int dyads, an unfortunate artefact of this material. Further, there is no basis here for comparing, between interviewer, the effect of variety of English, e.g. British, American, Australian, on the interviewee.

Of further interest is, as can be seen from the transcript of the dialogues, that each of the three Finnish interviewees had lived abroad at some period in her life. As the interviewees did not give any standardized biographical information, it is impossible to determine differences in time abroad or degree of time spent speaking English. However, one interviewee mentioned teaching in English and one conducting research at an American university.

All of the interviewees were adults, each a professional in, apparently, a different field. All were academically educated, as is revealed on the videotapes. This fits well with the aims of National Certificate testing. The tests are intended for adults in particular, and for anyone regardless of educational background. Naturally, at the advanced level in languages, most testees have academic backgrounds.

I have labeled the interviews 1, 2, and 3, according to the order of the videotapes when I removed them from the mailing envelope. There was no intention to create any bias in so doing.

The individuals in the dyads in the three interviews are labeled, named, in the following anonymous manner:

1	Maija Meikäläinen	John Doe
2	Kaija Karkkinen-Smith	Jack Doe
3	Liisa Lintunen	Harry Doe

These labels replace names on the videotapes, when given. On the NC, each interviewee is asked to give her name clearly, in full on the videotape. This is because the evaluation of testees in the NC is made on the basis of the videotapes. The interviewer may, after the interview is finished and the candidate has left the room, jot down a few impressions for the evaluator, but the evaluation is a separate process, done on the basis of the videotape. Even though I stated above that there were technical difficulties with some of the recordings, I must hasten to add that these were surely not grave enough to cause difficulties with the evaluation of these candidates. These technical difficulties mentioned did cause problems with transcription, simply because the level of recording necessary for fine transcription, such as that used in my study, is very demanding indeed. One need not hear absolutely every word from a videotape in order to evaluate a performance, but such clarity is necessary for fine transcription.

3.2 Methods

3.2.1 Transcription

Over two decades of work at the University of Helsinki Language Centre, I have transcribed several hours of tapes, running to hundreds of pages. I think it is fair to say I was familiar with the mechanics of transcription, even with some of the problematics of transcription. Nevertheless, in

approaching this study, I had to look into the literature for a “best usage” of transcription. It was no great shock to find that each study seemed to utilize a different transcription method, often depending on external factors such as type of computer program, whether or not IT data compilation and analysis were being used. More to the point, the studies attempted to tailor their transcription usage to fit the purpose at hand. Therefore, I did the same, attempting not to take any a priori assumptions about the outcome, but simply trying to include as much detail as could be taken in with a simple procedure.

Some transcription methods used in IT compilation and analysis cannot be described as transparent, as it is impossible to display all of the salient features at one time, at least in a printed journal or book. Therefore, I tried to develop a transcription method which is printable, transparent, straightforward, and, in particular, detailed enough to yield analysis of interaction (cf. White 1997).

3.2.2 Transcription conventions

The primary considerations are, first, that each line of text in the transcription corresponds to one c-unit. C-units (Brock 1986: in Foster 1998: 8) are: “independent utterances which provide referential or pragmatic meaning, i.e. utterances which are meaningful though not necessarily complete.” See also discussion in chapter 2: Background. Second, the transcript is a table in two columns, left: speaker identification; right: utterance, i.e. c-unit.

The following are the main symbols used:

S	subject (interviewer/interlocutor)
I	interviewee (elsewhere in paper, Int)

In order to conserve space in the transcription I have shortened Int to I, thinking that it would cause no confusion, as it would in the text.

- [overlapping utterance, marked at the point of overlap
-] second utterance latched to first, no overlap (if not marked, but line starts immediately to the left of the column, it means the same; in other cases: the pause is marked)
- = continuity of c-unit to another line, used either because the c-unit is too long to fit on one line of the table or because of back channeling interrupting the line, which continues despite the interruption, sometimes used to mark the continuation of dialogue by a speaker, without intending to mark the c-unit (this being the greatest weakness of the transcription system adopted)

There are several other symbols adopted to assist in marking sufficient detail to make the analysis. These are as follows:

- [] paralingual, primarily pausing or phatic communion (markers)
- [.2] pause: time given in tenths of a second, e.g. two-tenths of a second
- [breath intake] breath intake, primarily a pausing device
- [sub] subvocalization
- /- return to normal speech after subvocalization
- [???] unclear word, jumbled word
- ?word slightly unclear word, e.g. ?what's

[?]	rising intonation
[,]	continuing intonation
[glottal stop]	glottal stop
[chuckle]	chuckle
[laugh]	laugh
[sound of smthg]	sounds, marked, made by one S/Int
[?smthg]	comments on observations from the videotape—at right edge of column
{ word }	unusual, variant pronunciation of a
word	which could possibly be understood otherwise
mmmmhmmm	lengthening of sound indicated by letters
com'parable	marked, unusually stressed syllable
-	word cut off, e.g. <u>wh</u> - truncated from <u>what</u>
+	sound(s) added on afterward, e.g. <u>religion</u> + z

3.2.3 Issues considered

In chapter 2, Background, some of the foundations for studying interaction in dialogue, particularly in interviewer-interviewee dyads in interviews, are reported. It was difficult to decide on a particular approach to the organization of these issues. They do not lend themselves readily to any hierarchy. However, I decided to take the most straightforward approach I could determine, especially after realizing that the literature afforded no appropriate model(s) to follow. I have decided to move from those factors which reveal interaction taking place to those which outline what type of interaction has taken place. Following this, I concentrate on what in the interviewer's input has had effect on the interaction.

4 RESULTS AND DISCUSSION

In order to move from the simpler to the more complex in the results, I have taken the following basic order of the issues involved. After reporting the basic timing, the turns and c-units are studied. Then I look at the starting sequences, followed by lexical considerations. Conversation management is the next topic, and it is divided into several sections. Negotiation of meaning and accommodation are the final sections of the results. Because much material would have needed to be repeated, I have chosen to collapse the results and discussion section into one.

The following is a brief breakdown of the areas to be covered:

1. basic info: turns, c-units
2. starting sequences: the effects of pitching by S
3. lexical density: its relation to interaction
4. conversation management
 - 4.1 back channeling: its appearance with other interaction
 - 4.2 overlapping
 - 4.3 pausing: filled and unfilled pauses
 - 4.4 introspection: some examples throw light on this
 - 4.5 repairs and repetitions: show interaction
5. negotiation of meaning and accommodation

Naturally, in the discussion of these areas, comments in certain areas will be made which point at some of the other sections. Since the division of this nexus is intransigent, it is probably wiser to follow this plan to rather freely discuss these issues within the framework of the larger picture.

4.1. Basic info

The 3 interviews each involved a dyad; each had a Finnish female interviewee and a British male interviewer. Six people were involved.

Table 4.1 gives the timing of each interview. Although each was approximately ten minutes in length, there were slight differences.

Table 4.1A Timing of the interviews.

no.	time	seconds
1	11' 09"	669"
2	9' 59"	599"
3	12' 50"	770"
	sum	2038"

The videotaped portion of these interviews, at least one was cut off for some reason, total 33 minutes 58 seconds.

Table 4.2 shows the number of turns in each interview. A high number of turns can indicate many things, ranging from a large number of questions asked by a rather desperate interviewer who is receiving single word answers to a very involved, interactive discussion. Unlike some analyses of turn-taking (cf. White 1997), I have chosen to count as a turn the utterance by one speaker until it is interrupted or truncated in some fashion by the other speaker, e.g. back channeling which often occurs within the flow of speech. In such cases, the c-unit may possibly continue, e.g. after mhmm. I have decided to treat this as a turn. In transcribing the videotapes, I was forced to conclude that phatic communion indicated interaction, as a matter of fact that is how it is often described (cf. Stubbs 1983).

Table 4.1B Turns in the interviews.

no.	S	Int	S/Int
1	85	85	170
2	74	73	147
3	139	140	279
totals	298	198	596

I do not think it wise to comment very much on this, except to note that actually very many turns occurred. Perhaps this could be said to possibly, given the outcomes of other analyses, point vaguely in the direction of interaction, unless the interview consisted of a series of questions asked by S followed by yes or no answers by

Int. In actuality, no such thing happens in the dialogues as can be seen from the transcripts.

Another interesting feature is to look at how much time each spent talking during a turn. Table 4.3 presents the average time per turn.

Table 4.1C Average time per turn.

no.	S/Int
1	3.9"
2	4.1"
3	2.8"
all	3.4"

Already from this some summary information may be obtained. But first it must be pointed out that collapsing Ss and Ints has done injustice to the differences between them. Ints spent a lot more time speaking, as can be seen from the item analysis presented below (4.3). Since none of the Ss spent much time speaking, it is safe to surmise that the interviewee in no. 2 took much longer turns than did Int in 3, a rather much longer than in 1. The transcripts bear this out; thus the data confirms the conjecture aroused during work on the transcripts. Int 2 takes longer turns, sometimes holding the floor for lengthier periods. Nevertheless, on the whole it must be admitted that the turns are rather short. That would indicate, at least until further judgment is in, that there is quite a bit of S/Int interaction.

The next important basic consideration is the c-unit. As previously discussed (cf. Brock 1986), a c-unit is an independent utterance which provides referential or pragmatic meaning, an utterance which is meaningful although not necessarily complete. Table 4.1D gives the number of c-units.

Table 4.1D Number of c-units in the interviews.

no.	S	Int	S/Int
1	111	122	233
2	87	108	195
3	151	168	319
totals	349	398	747

As can be seen from this data, Int clearly produced more c-units than S. Many of the c-units, as mentioned above are phatic communion of the mhmm, yeah-type. Most of these occur in the interviewers' speech. An adjusted table taking this into consideration is given below.

Table 4.1E shows the amount of short phatic communion in each interview, and table 4.1F the adjusted number of c-units.

Table 4.1E Short phatic communion in interviews, as c-units.

no.	S	Int	S/Int
1	40	19	59
2	45	14	59
3	100	22	122
totals	185	55	240

From this table it can be clearly seen that the third interviewer used more short utterances, mhmm, than did the others, even counting that the third interview was longer.

Table 4.1F Adjusted number of c-units in the interviews, discounting phatic communion.

no.	S	Int	S/Int
1	71	103	174
2	42	94	136
3	51	146	197
totals	164	343	507

This table shows that the interviewees produced a great deal more speech than did the Ss. In that sense the test interview dyads are not

actual discussions between two equal partners. However, as White points out (1997), even in actual speech events, e.g. business negotiations, one partner may, because of perceived or implicit cultural assumptions or due to intercultural bumps, take the floor more than the other. In this case, Int is attempting to demonstrate her best speech sample, while S is trying to assist in eliciting that sample. That may or may not introduce bias into an interview, or interviews in general, due to ceiling effects or validity considerations. Below, the number of c-units produced by each is rendered as a percentage, further facilitating the breakdown of this data.

Table 4.1G Percentages of c-units produced by S and Int.

no.	S	Int	S/Int
1	40.8%	59.2%	100%
2	30.9%	69.1%	100%
3	25.9%	74.1%	100%
totals	32.3%	67.7%	100%

In conclusion, it can be said that about one-third of the c-turns were short, phatic utterances. Of the remaining c-units, about one-third were produced by S, while the bulk, two-thirds, were produced by Int.

Foster (1998) presents some interesting data on number of negotiated moves per c-unit, but the data is so inconclusive that it is difficult to make direct comparison. However, I will attempt to return to this issue in slightly different form in 4.5.

4.2 starting sequences

Ross (1992: 175) in rather a different discussion suggests that the interviewer makes a rather quick assessment of Int's ability or performance based on the opening sequence, i.e. a sort of instant judgment on what happens immediately at start. In one sense, this is desired, because if S pitches the dialogue too high or too low that can have ceiling or floor effects on the overall production. NC interviewers are trained to avoid this (Appendix 3), both

- S I've never heard that [.5] city needing an interpreter
before
- S or perhaps
- I [aaah
- S = I've just never heard of it
- I [,] well aaahh there was aahh very little Russians before ah
nineties in Finland
- S mm
- I they have all just came come [breath intake]
- S [yes

Int starts with an idiomatic phrase but I'm coping followed by S's if you don't mind m- asking what are you doing at the moment.

Picking up on the cue of coping, S pitches the dialogue at a fairly high level, and it is matched by a series of utterances which demonstrate that the level is met. Each time an interview begins, the interviewer is placed into this awkward position of determining from minimalistic cues, non-linguistic as well as opening utterance, perhaps phonemics, as to how to pitch the initiation of interaction.

After the initial turns by Int, S challenges this initial assumption by stating alright so so for example [.9] th- the hospital could call you up any time say come and help interpret. Once again, Int responds very appropriately. Indeed the give and take demonstrated in the above sample give one hope when thinking of problem of the artificiality of interviews. The situation may not be as bad as some researchers (Bachman 1988) seem to indicate (to be fair Bachman is not always in print or person quite as negative as I may have represented).

Not only in the example above but also in the dialogues in this study, Ss show remarkable skill in initiating dialogue and achieving a suitable pitch. They all demonstrate further checks of this pitch, e.g. in interview 2 just after the opening sequence S states:

- S and ah ask you can you make a definition of literacy [.2]
because I think actually what your opinions are [.6]
- I [mhm
- S = of literacy ah depend very much on what you understand of
it I mhm and especially in modern times it is very
interesting issue because the literacy has changed from the
novels [breath intake] of ?poesy towards ah poe- poetry
towards something which is more to do with also with
computers and the computer literacy [breath intake]

This tentatively posed question (with 8 hedges, many interpersonal), carefully placed into the stream of the dialogue results in what is perhaps the longest c-unit in the three interviews. The result is certainly positive. When S initiates an interview, the opening sequence can result in a positive or a negative effect on the entire interview. This gambit provides a fulcrum on which so much of the framework of the dialogue then balances. If positive, as in these three interviews, then the interview proceeds on a good note.

4.3 Lexical questions/lexical density

O'Loughlin (1995) found statistical significance for lexical density in candidate output over test format, task type, and interaction between the two. Nevertheless, O'Loughlin points out that the differences are small. Without recourse to the entire scheme used, it is impossible to achieve commensurable results. In particular the division into high-frequency and low-frequency items is outside the realms of such a non-IT study as this.

Lexical density refers to the concept pioneered by Ure (1971), i.e. the relative percentage of lexical or grammatical words in a text as a portion of the total words. Ure found that spoken texts generally had a lexical density of less than 40%, with speech involving dialogue less than 36%. In contrast written texts generally had a lexical density exceeding 40%. This number, 40, then, became the watershed between spoken and written texts. Naturally, there was

some overlap in the range of the figures presented, but the general trends divided neatly at forty.

Halliday (1985) sees grammatical items (G) as function words, functioning in closed, finite systems in the language. Lexical items (L), conversely, are content words, in infinitely extendable open sets. The concept of set is rather clearly borrowed from mathematics in this sense. Halliday classifies modal and auxiliary forms of verbs as well as all forms of the verbs “to be” and “to have”, all proforms, interrogative and negative adverbs, etc. as grammatical items (in addition to words traditionally classified as grammatical, e.g. articles). O’Loughlin, arguing that the main consideration of a system of calculating lexical density is consistency in classification, extends this set above to include discourse markers, interjections, reactive tokens and filled pauses (mhmm).

Since I do not consider that the final word is in on this classification issue, I took the liberty to count also “to be” and “to have” where they had semantically meaningful function, e.g. I have a book contra I’ve been going there for three years. O’Loughlin counts multiword verbs and phrasal verbs (find out about or put off) as single items, which seems reasonable to me, on the basis that it forms one functional whole. On the other hand, contractions (haven’t) are counted as two grammatical items, which does not really seem quite justified. There is no argument introduced, so I would like to defend my usage by stating that contractions are used by the speaker as single units, not as multiplicities. Perhaps a bit more of phonemics should be introduced into the discussion of oral phenomena in general.

It could be argued that discourse markers, particularly those sometimes labeled particles, should not be counted in a “word” count. But when one is thinking of functional phonemic phenomenon in spoken language, I can think of no adequate reason for discounting them. Actually, I would argue that many studies are biased because they are made on the basis of the transcription of taped materials, not videotaped. All paralinguistic gesturing,

etc. is then lost. When I think back on it, I am appalled at my lack of including more comments on this particular aspect in my own transcripts. However, at one point in the transcript where a nod was used as a mark of affirmation, I did count that as a discrete grammatical item. It is functionally equivalent to a -judgmental reinforcer (see discussion below). This particular item is found in the following excerpt from interview 3:

S mhmm

I and they were Germans and Finns and Swedes living in same village

S [nod]

I and we sort of lived in our own [.8] community

S oh right

My findings were, generally, consonant with earlier findings as to lexical density. Table 4.3A gives the item counts for the three interviews.

Table 4.3A Item counts on the interviews, itemized for S and Int separately.

	S			Int		
no.	L	G	L + G	L	G	L + G
1	202	411	613	334	667	1001
2	166	336	502	527	818	1345
3	166	307	473	606	1004	1610
tot.	534	1054	1588	1467	2489	3956

This table shows that almost four thousand items were produced in the three interviews, the bulk of them by the interviewees. That is as it should be. S is attempting to elicit as good a language sample as possible, and relative amount of production must count in that. Collapsing the information in this table, we get the item total for each interview.

Table 4.3B Item counts on the interviews, S and Int combined.

	S + Int		
no.	L	G	L + G
1	536	1078	1614
2	693	1154	1847
3	772	1311	2083
tot.	2001	3544	5544

The lexical items comprise only some 36% of the total. Next, I would like to look further into the relations between the lexical items and the grammatical items, in order to allow for a bit of comparison with O'Loughlin (1995). Table 4.3C gives these percentages.

Table 4.3C Percentages of lexical and grammatical items on the interviews, itemized for S and Int separately.

	S			Int		
no.	L	G	L + G	L	G	L + G
1	33.0%	67.0%	100%	33.4%	66.6%	100%
2	33.1%	66.9%	100%	39.2%	60.8%	100%
3	35.1%	66.9%	100%	37.6%	62.4%	100%
avg	33.6%	66.4%		37.1%	62.9%	

These figures are very consistent indeed. And they are certainly within the parameters described by O'Loughlin (1995), i.e. less than 40% for spoken language. Indeed, it is fair to say that two-thirds of the dialogue are made of grammatical elements, indicating a rather high degree of social interaction in the interviews. O'Loughlin (1990: 230) claims that more social interaction leads to lower lexical density. From the very slightly elevated percentages on L for S on the third interview, one might be tempted to surmise that that interviewer asked more content questions. As will be seen below, that is not exactly the case. The Int who produced the most lexical items, interviewee two, was the one who produced more content.

Table 4.3D demonstrates the relative item output by interviewees and interviewers.

Table 4.3B Relative item output on the interviews, S and Int.

no.	S	Int	S + Int
1	38.0%	62.0%	100%
2	27.2%	72.8%	100%
3	22.7%	77.3%	100%
avg	28.6%	71.4%	

From this table, it is easy to see that the vast bulk of the output was by the interviewees. Nevertheless, there is no such great dearth of items on the interviewers' side which would lead one to believe that this was a monologue. A perfunctory glance at any point in the tapescripts (Appendix 1) would demonstrate that there is much interaction taking place but on the terms of Int.

4.4. Conversation management

In conversation management, I want to look at back channeling, overlapping, pausing, introspection, as well as repairs and repetitions. All of these are aspects of conversation management and strategies, as demonstrated in the Background.

4.4.1 Back channeling

Following the lead of White (1997: 323), I have used his back channeling taxonomy in an analysis of these interviews. The basic framework calls for a division of back channeling into -judgmental, reinforcers and claimers, and +judgmental, prompters and clarifiers. All of these have to do with maintaining and claiming floor in the interaction. A good number of them in a dialogue should be a fair indication that some degree of interaction has actually taken place. Table 4.1A gives the number of each of these categories on the three interviews.

Table 4.4.1 Amount of back channeling, divided between +judgmental (+J) and -judgmental (-J), differentiated for S and Int.

	S		Int		
no.	+J	-J	+J	-J	totals
1	44	17	10	6	77
2	38	18	5	11	72
3	43	81	13	6	143
tot.	125	116	28	23	292

As can be clearly seen from this table, quite a lot of back channeling was used. So, it could be surmised that, once again, quite a lot of interaction occurred. Stansfield (1991) states that OPI and SOPI correlate to such a high degree largely because neither allows for much candidate interaction. Perhaps, this was an artefact of the material in that study. At least here it would seem quite clear that back channeling is found more from S than from Int, but Int is also demonstrating interactive ability. Perhaps an example from interview 3 would help to demonstrate this (back channeling underlined):

- I when we goes to these mass [.6] happenings [breath intake]
 S mhmm
 I = like concerts [1.1] or ah [.6] festivals somewhere [1.7]
 S yea but surely when most people go abroad they want to see something [.2] different
 S let me give you an example
 I [mhmm]
 S = if you were to go to London to see British soldiers in
 I [mhm]
 S = big red tunics with big high hats
 I [mhmm]
 S = marching [1.1] [breath intake]
 I [mm]
 S = ninety-nine percent of the British soldiers aren't like that
 I right, yeah
 S [but they go to see the difference [1.4]
 I [yes]
 S = not the normal ordinary people

- I eah [.2] I think that's ah {phenomenen} which is it's actually
 I we have the TV [.7] and the books [.9] which add to
 tourism [.8]
 S [mhmm

When S begins fronting this theoretical question to Int in the middle of a series of utterances from Int, Int moves smoothly into back channeling, demonstrating both +J and -J moves. This looks so much like dialogue between two rather equal partners that it would be really difficult to establish that this was not actually two friends speaking. The floor moves very smoothly from one speaker to the other.

Interviewer 3 used more back channeling than anyone else. Most of this consisted of the phatic mhmm or yeah. In a final note, I would add that on interview 1 a chuckle was counted as a +J prompter. I have not found such a classification in the literature, but it seemed to fit the functional description rather well.

4.4.2 Overlapping

Overlapping, where one speaker speaks on top of another's speech is an interesting phenomenon, particularly since Halmari (1993) found that Finns in intercultural telephone conversations overlapped the last element rather than in the middle of the interactor's turn. Table 4.4.2 gives the number of overlaps mid-utterance (M) and last element (L). From this description, then, it should be expected that the British S overlaps relatively more in mid-utterance than does the Finnish Int.

Table 4.4.2 Number of overlaps on each interview, mid-utterance (M) and last element (L), differentiating between S and Int.

	S		Int		
no.	M	L	M	L	totals
1	26	5	5	3	39
2	28	5	9	7	49
3	69	9	11	9	98
tot.	123	19	25	19	186

Although this sample is too small for statistical analysis, it clearly demonstrates a trend. Finnish Ints used mid-utterance overlapping more than they did last element overlapping. Furthermore, in interview 1, all the last element overlapping was in answer to direct questions. In interview 2, all the last element overlapping was to answer question or to follow on prompts. The situation on interview 3 was the same as on interview 2. I do think that Finns commonly use mid-utterance overlapping at least those whose English is adequate to participate in the Advanced level of the National Certificates testing.

Looking back at the interviewers for a moment, there is a large discrepancy between S3 and S1 and 2. Even though interview 3 is somewhat longer, that is not enough to explain the difference. Once again, the difference is found in the great amount of phatic communion of the mhmm type. As a trainer of test interviewers, I would probably have noted to S3 that he probably challenged Int more, not relied so much on simply noting that he was paying attention to Int's speech. However, not being in the test situation myself, I may be reading too much into this. Int 3 could see the facial reactions of the candidate in a way which did not show on the videotape and, thus, may have been receiving clues or signals that this was the correct approach with this candidate. Certainly, it must be admitted that the language sample seems quite representative and rather uniform throughout, not changing pitch when S did prompt challenges.

4.4.3 Pausing

These dialogues are ridden with pauses, both filled and unfilled. The most common filled pause is the breath intake, obviously used to hold floor, gaining time for thought, searching for the right term, or indicating willingness to relinquish floor. Examples of these can be found in interview 2:

- I [yes and if you would want to measure ah literacy as a certain critical eye [breath intake] as ah something that when you are literate you also need to be [.6] ah critical towards what is ah in the text
- I that is very difficult to measure because that's more kind of like a cultural literacy
- I how are you able to understand the message [breath intake] the symbols the hidden meanings [breath intake]
- S [mhmm
- I = how to measure this
- I but of course it is important because that's an important part of our culture is this kind of hidden [1.4]
- S [yeah
- I = levels
- S yeah [.8] are these themes or topics that come out from your own research or [.6] in a sense that [.2] ah presumably writing in medieval times [.2] there was the same [.9] an audience [.2] an author [.2] and to some extent you know

An example showing unfilled pauses in an interviewee could be taken, e.g. from interview 3:

- I [breath intake] ah [.7] so we think more that we can share
- I there's a lot to learn from the community and society life [.5] and social [.7]
- I what we have now here [1.1] established
- S [mhmm
- I = new things you know
- I you will have to learn work gr-
- S [mhmm

- I = working in a group [.5] and take care of each other [.9] and
pay attention to your
= neighbor and so on
- I these things [.6] are social matters
- I I think we w- would have a lot to learn from them
- S mhmm so I mean does [.4] traveling and living abroad [.7]
actually help [.3] global understanding

It may be that NSs use more unfilled pauses, particularly between tone groups, but it seems that NNSs also use them quite often. The unfilled pauses above seem to be basically of the type: searching for content (what), not: searching for how to express something already formed in L1 (how).

Jefferson (1989) makes quite a bit of manually timing pauses down to the hundredth of a second and then using this for statistical analysis, attempting to determine which length of pause is used for what purpose. To me, this seems to be going a bit far. First, you cannot simply time accurately, using a hand-held stopwatch down to .01 second. It is difficult enough to get it to the tenth of a second. I had to time each pause at least twice, whenever I got the same result, or three times (then I took the average). Pauses, etc. are examples of features which should be subjected to IT analysis.

4.4.4 Introspection

Private speech is also reported to be one of the factors in interaction (White 1997). One example could be taken from interview 1:

- I aahh they were called [.8]
= [aside] ah mm [smacks lips] I don't know how do you say it
in English
= but ah like ah Inkeri [2.4] person [.2] people in Inkeri land
- I and aahh
- S [mmm
- I = they were let to come in Finland in the nineties

Int is reporting on the lack of an English equivalent for Inkeri. The Latinate Ingria is something this candidate may not have previously encountered, prompting this bit of introspection.

Another example can be found in interview 3:

- I = who behave differently
 I [breath intake] so [.8] I think it's a matter on a personal level
 S mhmm
 I = ahm [.8] I had an idea I forgot [.2] what was it [2.3]
 I ah just continue [.2] i- if it comes up I'll tell [giggle]
 S [okay wh-
 what about this concept
 S people sometimes talk of a global village

In this instance, S makes a play and recovers the situation for the partner by proposing a new line of questioning on the topic under discussion.

I did not uncover any instances of Int focusing on the test assessment during the test. In some way, it would appear that the Ints were probably focusing on content rather than on process, the process, interaction, that came out did so without any particular effort on their part.

4.4.5 Repairs and repetitions

There were many examples of repairs and repetitions in the interviews. Following White (1997), I looked not only for straightforward repairs and repetitions, but for tentativeness (in both S and Int, particularly in S), and modality in verbs. All these indicate a close interaction between the two participants of the dyad. An example of repetition can be found in interview 2:

- S dare I ask you what is your area
 S I was looking [gestures to papers]
 I [it's medieval history
 S [medieval hist-
 I [= history of
medieval ah ah religion +z -gions and especially that of lay
 laywomen
 S mhm
 I so I'm studying the religious ah life of a medieval lady

This exchange establishes not only the field of study of Int but also serves as a sociocultural exchange between S/Int, so that the interview, where S started out looking a bit nervous, seemed to turn a corner at this juncture. The two began to treat each other as equal partners in building the dialogue.

Repetition may have other functions, as in this example from interview 1:

- I and ah not all the people know Russian anymore [.8] like it was in the Soviet Union
 S so any- anyway th- be- from what I know they are a very proud nation
 I yeow
 S and now they've lost a lot of that
 I yeow
 S [that pride
 I yeow yes I I don't know what's the problem because ahh [breath intake] ahh many of them are are very well educated
 I and [.7] they know their profession
 I they could work here easily
 I but [.4] Finnish language is [.7] the problem
 S yeow yeow

In this example, repetition by S serves as a prompt, and apparently, judging from the outcome, the prompt is fairly successful.

Another instance of the sociocultural construction of meaning can be found in interview 2:

- S [yeow [.2] I was thinking th- that it's maybe easier to measure
= [sub] ?? /one of the old-fashioned definitions of it cause see
[.3] what's
= the difference the more modern the more
- I [yes and if you
would want to measure ah literacy as a ce- certain critical eye
[breath intake] as ah something that when
= you are literate you also need to be [.6] ah critical towards
what is ah in the text
- I that is very difficult to measure because that's more kind of
like a cultural literacy
- I how are you able to understand the message [breath intake]
the symbols the hidden meanings [breath intake]
- S [mhmm
- I = how to measure this

Thompson (2001) argues that modality and reference to the “other” in text, as well as disclaimer and hedges, have mitigating effects, that they are moves within a text to bring the other into the interaction. If that is so of written text it must be even more so of spoken face-to-face interaction. This, then, demonstrates a high degree of interaction. The use of ah in the text above also serves an interactive function, underlining what is to come, i.e. ideas the speaker wants to emphasize.

Typical repairs are such items as when Int realizes that something has not come out quite right. The strategy then taken is interesting, as in this example from interview 1:

- I [laugh] ah of course ah but the time was {diffent} then
- S [yes
- I there was open {wacancies} in in Finland
- S [uhmm
- I [.] a lot of jobs
- S yes

This self-repair went quite smoothly. Surely S understood the utterance and did not call for any repair, i.e. a -J reinforcer. The continuing intonation marked for Int shows that she was determined to communicate, and this was rewarded by a +J reinforcer from S.

Another example of self-repair can be found in interview 3:

- I = th- the biggest problem perhaps
 S [uhuh
 I = that if we lived [1.8] as they lived there might be [.5] some kind of a way [.8] to go
 I but I figure that [.5] sometimes the color [.6]
 I
 S [mhm
 I = we didn't see that they are Blacks
 I and we were Whites

Apparently, Int replaces out with to faze out (impressionistic spelling) or to phase out. This is corrected immediately without pause, reflecting a rather good strategic use of language.

An example where Int is challenged on a concept and makes a repair is found in interview 3:

- I but there's always a danger th- that it [.6] makes us more [.5] {prejudiced}
 S [mhmm
 I = and strongly [1.2] ah to op- [.8] to- not to oppose [.7] but to
 S [?]
 I [[.5]
 S [right
 I = [1.2]

This challenge, or probe, by S proves very effective in affecting a repair and actually negotiates the meaning of an unclear concept. Notice the +J reinforcer provided by S upon mutual understanding.

From the examples above, it can be seen that conversation management strategies by repetition and repair worked fairly well in these interviews.

4.5 Negotiation of meaning and accommodation

Accommodation appears to subsume both itself and negotiation of meaning. While there is not much that could be on the micro-level stated as negotiation of meaning in these tapescripts, yet on the overall language interview-as-text level these interviews are largely the production of and forwarding of a common understanding, i.e. the negotiation of meaning.

Interview 1 has a good example of negotiation of meaning in a clarification question by Int:

- S] right [.2] but [.3] wouldn't you say
 S that there was a world of difference between say [.2] me who comes to Finland [.9]
 I [mhm
 S and [.9] unless people speak to me I could pass in the street for a Finn [breath intake]
 S and you going to Tanzania
 S and [1.0] being an outsider who's come temporarily
 S very clearly not a local
 I [mmmmm [1.1] th- there's a big difference [.9]
 I did you ask that
 S [ah that's right
 I = yes [.5] [breath intake}

This is not an introspective question but a question directed to the partner, requesting clarification, which then after a +J reinforcer I begins to answer.

Ross (1992) has seven categories of accommodation used for classifying interview questions. I only found the first three of them in these interviews: display question (D), or-questions (O), and fronting (F). I did not find the following: grammatical simplification, slowdown, over-articulation, and lexical simplification. Naturally, as can be seen, these categories divide easily into two groups, one for use with beginners and the other with, actually, everyone. Since the NC candidates on these interviews were at the Advanced level, the latter group did not come into question. Table 4.5 gives the amount of each type of accommodation question.

Table 4.5 Number of accommodation questions by category (D, O, F) on each interview.

	D	O	F	sum
1	20	7	6	33
2	15	7	3	25
3	11	12	7	30
	46	26	16	88

The first thing to note is that all of these accommodation questions received appropriate responses. The curious note is that the longest interview, #3, had fewer display questions than did the other two interviews. This is an outcome of the strategy employed by S. Actually, it rather surprised me that this interview was upheld not so much by questions and answers and forms of accommodation and negotiation of meaning but by phatic communion (social skills). Perhaps this interviewer was more “on the level” of the Int as an equal partner. Perhaps it is only the outcome of an interviewing strategy. Other anomalies in the figures for this interview were noted above.

Ross’s (1992) transcribed examples are at a Basic level, not Advanced, so they are not directly comparable to my data, and this

factor may have skewed the outcome, one way or the other. Nevertheless, the important factor to note is that quite a few accommodation questions were used in my material and that these received appropriate responses.

I think that a good example of the negotiation of meaning is found in interview 1:

- I [and of course very important to know all the manners [.3]
and
I so that you you don't make all the horrible mistakes
S [uhm
I = at first [swallows] first day
S from your experience do you think that knowing the dos and
don'ts of a culture [.9]
= can actually help you or hinder you [.8] when working in
that culture
I [1.1] you mean ahh how in what way {hiinder}
S ahh [1.4] they can be [.7] by knowing them you can be put a-at
a disadvantage
I [2.3] ah well I think it helps you [rolls eyes] [.7] if you know
[.2] know cultural differences
I although maybe you can't [1.1] ah can't follow them [.8] in
every occasion
I you have to be ah neutral [1.2]
S right
I ah you can't take a side in every situation [?? any]
S right

Both partners take several turns in this negotiation. First S leads by offering options, and then they begin to negotiate the meaning of an intercultural theme. Then S gives +J reinforcers to indicate that common meaning has been achieved.

I found an interesting example of Int treating S as an equal conversation partner in interview 3:

- I = and [.2] there is [.6] there are many gaps ge- to cross over
 S] is it a question of money as well
 I [aw that's the main question
 I = th- the biggest problem perhaps
 S [uhuh

This is from the view of checking S interaction and S's effect on interaction a very reassuring bit of data. This actually elicits a line of discussion which continues for many pages, indeed remains in some form to the end of this interview (it is about one-third of the way into the interview).

Since turn-about is fair play, I looked for and found an example of Int accommodating S. This example comes from interview 1:

- I [yes it was I think it was a [.3] a mistake
 S yeow [.4] yeow [.5] what about these ah
 S [1.5] ah uh how could I explain
 S = / the nouveau riche
 S could I translate these as the [.2] Moscow [click] doing all
 their shopping in Helsinki
 S mhmm
 I has this had a great impact on [.7] Finland 'd you say
 I [.8] ahhh [.5] well I think uh [.3] well in in [1.0] uhmm
 [smacks lips] eastern parts of Finland
 I I lived there for four months [.7] on winter [1.0] last winter
 I and uhm [.8] I think it's [.3] ah they are quite [1.3]
 I ah well [.9] you can see them very much ah [1.7] in the cities
 I and they are buying a lots of stuff
 I and there are lots of new shops [1.5] ahh for them [.7]
 S [uhmh
 I = especially for them

In this example, I do not think we can be quite certain that S was unable to make up a question. Perhaps this was a question being formatted very tentatively. Or perhaps S was, after all, searching for a question, for a concept. S certainly gives good signals to the partner, e.g. the modals (could), the phatic communion (mhmm),

but particularly the click of the pen. Interviewers have even been warned not to hold pens in their hands when giving interviews (it might give the impression that notes are being taken or assessment marks awarded). But in this case, the click certainly serves as a speech act, probably as an act indicating willingness to relinquish floor.

It is very interesting in this example that Int proposes her own question. This is then discussed for several turns. This action can surely be interpreted as a negotiation for meaning.

In the next brief chapter, then, I will try to ponder some of the implications of these results. These three interviews provided a wealth of material for study.

In the praxis of OPI, several issues continuously come under the scrutiny of research. Study ranges from major questions to specific, technical details. Some of the broad areas of study would be the ethics of testing and intercultural bias. Specific research categories could be represented by pausology and back channeling.

An ethical OPI should, for example, be able to document that the testing situation itself does not intimidate the candidates. One way to demonstrate non-intimidation is to show free interaction. An interculturally sensitive OPI should give evidence of having adequately taken into consideration sociocultural matters. The extent to which an OPI manages this can be illustrated by reference to interaction. The examples of specific details, pausology and back channeling, are held to be factors in interaction. It is safe to say that interaction is metaconcept present in some form or another in almost all of OPI. It is a major indicator and descriptor. Thus, if it can be shown that active interaction has occurred, several testing requirements have been met.

In my study I believe I have shown that the oral proficiency interview in the National Certificates testing resulted in a wealth of interaction, both in amount of interaction and in terms of the broader range of specific factors involved. I demonstrated not only that turn taking in the dialogue shifted about every third second but also that these turns varied greatly in length both for S and Int. The c-unit analysis began to reveal that the interviewee had the larger proportion of the dialogue. These units were shown in examples given under accommodation not to be a burden or imposition on Int but the outcome of interaction. Further, short phatic communion was an extremely common feature, once again for both partners in the dyad. The opening sequences gave evidence of the setting of appropriate pitch to meet Int's level. Pitch was then checked at later points in the interviews.

The analysis of lexical density yielded some interesting information. First, lexical items accounted for around 35% of the total, thereby demonstrating a rather high degree of interaction.

Second, Int produced well in excess of two-thirds of all items. This shows that the interaction that is taking place is happening on Int's terms. Back channeling was used by both partners in the dialogue. Both S and Int used more +judgmental back channeling, mostly reinforcers, than they did -judgmental. This enhances interaction. The study of overlapping gave a surprise. Earlier studies with Finns had found a prevalence of last element overlapping over mid-utterance overlapping. The results on my study were quite the opposite. Mid-utterance overlapping is generally held to be an indication of greater interaction than is final element overlapping. Interaction was also shown in the prevalence of filled and unfilled pauses, often indicating the maintaining and shifting of floor. Private speech and introspection showed that the candidate felt secure enough in the interactive framework to comment on her own oral production. The data evidenced a variety of functions of repetition and repair, particularly self-repair.

I also gave multiple examples of accommodation and the negotiation of meaning. The count of accommodation questions revealed that there were more display questions. However, the longest interview (3) relied less on display questions and more on phatic communion. That was quite an interesting finding, emphasizing the sociocultural nature of these interviews.

At various points in my study, I noted that the results of particular studies were not directly comparable to my data. This was often the result of the level or age of the candidates in those studies. Research reports on, for example, basic level candidates in an OPI give completely different results. Indeed, one can only wonder why OPI was administered at this level.

In thinking of what implications my study may have for further research, I would say that the instruments used in studies of oral language need to be refined. The c-unit is a good step forward, but a functional definition should be arrived at which would clearly divide the data. Similar comments could be made about the other tools of the trade, particularly the various taxonomies. Nevertheless, it must be admitted that much can be analyzed from oral usage with these tools. Yet, somehow the more I look into oral language not only do I understand more but also I further sense the mystery of language and sociocultural interaction.

BIBLIOGRAPHY

Bachman, L. 1988. "Problems in examining the validity of the ACTFL oral proficiency interview." *Studies in Second Language Acquisition* 10: 149-64.

Bachman, L. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Barnwell, D. 1989. "Proficiency and the native speaker." *ADFL Bulletin* 20, 42-46.

Brock, C. 1986. "The effects of referential questions on ESL classroom discourse." *TESOL Quarterly* 10: 47-59.

Brown, P. and S. Levinson. 1987. *Politeness: Some Universals in Language Use*. Cambridge: Cambridge University Press.

Chaudron, C. 1988. *Second Language Classrooms*. Cambridge: Cambridge University Press.

Chomsky, N. 1975. *Reflections on Language*. London: Temple Smith.

Davies, A. 1989. "Communicative Competence as Language Use." *Applied Linguistics* 10/2, 157-170.

Foster, P. 1998. "A classroom perspective on the negotiation of meaning." *Applied Linguistics* 19/1:1-23.

Goldman, A. 1994 "Doing business with the Japanese: A guide to successful communication, management , and diplomacy" in D. D. Cahn (ed.) 1994: *SUNY Series in Speech Communication*. Albany, NY: State University of New York Press.

Graham, J. L. 1985. "The influence of culture on the process of business negotiations: an exploratory study." *Journal of International Business Studies* (2nd quarter): 29/94.

Hall, J.K. 1993. "The role of oral practices in the accomplishment of our everyday lives: the sociocultural dimension of interaction with implications for the learning of another language." *Applied Linguistics* 14/2.

Halliday, M.A.K. 1985. *Spoken and written language*. Melbourne: Deakin University Press.

Halmari, H. 1993. "Intercultural business telephone conversations: A case of Finns v. Anglo-Americans." *Applied Linguistics* 14/4:408-30.

Hamilton, J. M. Lopes, T. McNamara and E. Sheridan. 1993. "Rating scales and native speaker performance on a communicatively oriented EAP test." *Language Testing* 10/3.

Hasan, A. A. 1988. "Variation in Spoken Discourse in and Beyond the EFL Classroom." Unpublished Ph.D. thesis, University of Aston. Reported in: O' Loughlin, K. 1995. "Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test." *Language Testing* 12/2: 223.

Hayashi, T. and R. Hayashi. 1991. "Back channel or main channel: A cognitive approach based on floor and speech acts." *Pragmatics and Language Learning* 2: 119-38.

Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Hunt, K. W. 1966. "Recent measures in syntactic development." *Elementary English* 43: 732-9.

Jefferson, G. 1989. "Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation" in D. Roger and P. Pull (eds.). 1989. *Conversation: an Interdisciplinary Perspective*. Clevedon and Philadelphia: Multilingual Matters.

Lantolf, J. P. and G. Appel. 1994. "Theoretical framework: An introduction to Vygotskian approaches to second language research" in J. P. Lantolf and G. Appel (eds.) *Vygotskian Approaches to Second Language Research*. Norwood, NJ: Ablex.

Lantolf, J. P. and W. Frawley. 1985. "Oral-proficiency testing: A critical analysis." *Modern Language Journal* 69: 337-45.

Lantolf, J. P. and W. Frawley. 1988. "Proficiency: understanding the construct." *Studies in Second Language Acquisition* 10: 181-96.

LoCastro, V. 1987. "'Aizuchi': A Japanese conversational routine" in L. Smith (ed.) 1987: *Discourse Across Cultures: Strategies in World Englishes*. New York: Prentice Hall.

March, R. B. 1989. *The Japanese Negotiator: Subtlety and Strategy Beyond Western Logic*. Tokyo: Kodansha International.

McCafferty, S. G. 1994. "The Use of Private Speech by Adult ESL Learners at Different Levels of Proficiency" in J. P. Lantolf and G. Appel (eds.) *Vygotskian Approaches to Second Language Research*. Norwood, NJ: Ablex.

McCreary, D. R. 1986 *Japanese—U.S. Business Negotiations, Praeger Special Studies. Praeger Scientific*. New York: Praeger.

North, B. 2000. *The Development of a Common Framework Scale of Language Proficiency*. Theoretical Studies in Second Language Acquisition vol. 8. New York: Peter Lang.

O' Loughlin, Kieran. 1995. "Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test." *Language Testing* 12/2: 217-37.

Pica, T. 1987. "Second-Language Acquisition, Social Interaction, and the Classroom." *Applied Linguistics* 8/1: 3-21.

Pica, T., L. Halliday, N. Lewis, and L. Morgenthaler. 1989. "Comprehensible output as an outcome of linguistic demands on the learner." *Studies in Second Language Acquisition* 11: 63-90.

Roach, J. O. 1945. "Some problems of oral examinations in modern languages: An experimental approach based on the Cambridge Examinations in English for foreign students, being a report circulated to oral examiners and local examiners for those examinations." Cambridge: Local Examinations Syndicate. In: Spolsky, B. 1990. "Oral examinations: An historical note." *Language Testing* 7/2: 158-73.

Rose, K. 1994. "On the validity of discourse completion tests in non-Western contexts." *Applied Linguistics* 15/1.

Ross, S. 1992. "Accommodative questions in oral proficiency interviews." *Language Testing* 9/2: 173-86.

Ross, S. and R. Berwick. 1990. "The discourse of accommodation in oral proficiency interviews." Paper presented at the RELC Seminar in Program Evaluation and Language Testing, Singapore, April 9-12, 1990.

Schegloff, E., G. Jefferson, and H. Sacks. 1977. "The preference for self-correction in the organization of repair in conversation." *Language* 53:361-82.

Schwartz, J. 1980. "The negotiation for meaning: repair in conversations between second language learners of English" in D. Larsen-Freeman (ed.) 1980: *Discourse Analysis in Second Language Research*. Rowley, MA: Newbury House.

Sengupta, S. 1998. "Peer evaluation: 'I am not the teacher'." *ELT Journal* 52/1: 19-28.

Shohamy, E. 1994. "The validity of direct versus semi-direct oral tests." *Language Testing* 11: 99-123.

Sindermann, G. and M. Horsella. 1989. "Strategy markers in writing." *Applied Linguistics* 10/4.

Spolsky, B. 1990. "Oral examinations: An historical note." *Language Testing* 7/2: 158-73.

Stansfield, C. 1991. "A comparative analysis of simulated and direct oral proficiency interviews." In Anivan, S. (ed.): *Current Developments in Language Testing*. Singapore: RELC, 199-209.

Strevens, P. Comments on Bachman, L. *et al.* *Final report of the Cambridge-TOEFL Comparability Study*. Mimeo. In: Spolsky, B. 1990. "Oral examinations: An historical note." *Language Testing* 7/2: 158-73.

Stubbs, M. 1983. *Discourse Analysis: the sociolinguistic analysis of natural language*. Oxford: Basil Blackwell.

Stubbs, M. 1986. "Lexical density: A computational technique and some findings." In Coultard, M. (ed.): *Talking about Text*. Birmingham: English Language Research, University of Birmingham, 27-48.

Swain, M. 1996. "Three functions of output in second language learning" in G. Cook and B. Seidlhofer (eds.): *Principle and Practice in Applied Linguistics*. Oxford: Oxford University Press.

Ulijn, J. M. and T. R. Gorter. 1989 "Language, culture and technical-commercial negotiating" in H. Coleman (ed.) 1989: *Working with Language*. Berlin and New York: Mouton de Gruyter.

Ure, J. 1971. "Lexical density and register differentiation." In Perren, G. E. and J. L. M. Trim (eds.): *Applications of Linguistics*. Cambridge: Cambridge University Press, 443-52.

White, R. 1997. "Back channelling, repair, pausing, and private speech." *Applied Linguistics* 18/3: 314-344.

Zora, S. and C. Johns-Lewis. 1989. "Lexical density in interviews and conversation." *York Papers in Linguistics* 14, 89-100.

APPENDIX 1

TAPESCRIPTS OF THE INTERVIEWS

Tapescript 1	73
Tapescript 2	86
Tapescript 3	100

INTERVIEW 1

dyad identified as: Maija Meikäläinen (I) and John Doe (S)

Impressionistic notes on the videotape:

[interpreter of Russian for the city of Vantaa]

[S questions quite factual, rather flat voice right up to the end]

[1. interlocutor off camera; 2. big chair in messy teaching room]

[Int uses quite flat intonation, and so does S (in contrast to his intonation on the previous interview—which could not be used for lack of permission)]

tapescript:

S [camera mike comes on mid-sentence] ... explodes it's my fault really

S just th- the [.7] sticking the video on [.8] [breath intake]

S = and [3.7] [clicking of buttons on tape recorder]

S the tape's there [.5]

S now you can forget about it

I alright

S okay [.3] nn just for the [.7] record could you state your name please

I my name is Maija Meikäläinen

S Meikäläinen's wonderful name to pronounce
 I [thank you
 S = by foreigners Meik- Meikäläinen
 I [yeow
 S okay [breath intake] how's it going
 I [breath intake] well it's been quite difficult [.8] for this time [sub] be-
 I but I'm coping
 S [good
 I = I hope
 S [that's the main thing
 S [breath intake] and could you actually if you don't mind m- asking
 = what are you doing at the moment
 I well I graduated this summer
 I ah my major was ah or it's Russian language and
 S [ahah
 I = literature
 I and now I'm working as an interpreter for the city of Vantaa [breath intake]
 S ahah
 I I go everywhere where there are Russian people Russian immigrants
 I and they need ah assistance [.6] for maybe social security or
 S [aa

- I = at health centers or
- S [alright so so for example [.9] th- the hospital could call you up
= any time
- I [yes
- S = say come and help interpret
- I yes so the city of Vantaa sends me to [breath intake] hospital of [.5] Meilahti
= for example
- S ahah alright is this a new position
- S I've never heard that [.5] city needing an interpreter before
- S or perhaps
- I [aaah
- S = I've just never heard of it
- I [,] well aaahh there was aahh very little Russians before ah nineties in Finland
- S mm
- I they have all just came come [breath intake]
- S [yes
- I = in ah nineteen [.7] ninety- two or ninety-five or [.3] between these years
- S yeow yeow
- I that's why this position is also quite new
- S how many Russians are there living [.6] in Finland now
- I well I know only th- the area of Helsinki

I it's more than ten {tousand}
 S [wal
 I = about fifteen thousand I think
 S yeow [.7] the whole of Finland you can double that easily I suppose
 I well I think aah most of them live here
 S [ahah okay
 I [= in this area [1.2]
 S ahah okay [1.5] right [breath intake] so it's [1.6] a huge [.3] percentage of the
 = foreign population [1.1]
 I [yes
 S = must be
 I yes uuhh I think it's [.3] it's one of the biggest
 S yeow
 I I'm not quite sure
 S yes [.7]
 S why are they coming to Finland
 I well uuhh [.8] [sub] I don't know if you /you probably have heard
 I but I can tell you anyway [breath intake]
 I ah when President Koivisto was [.7] on on [.3] was governing [breath intake] [1.2]
 I he decided that people who had lived in the area o- of Finland [breath intake]
 = in ah seventeen hundred [1.2]
 I aahh they were called [.8]

- I = [aside] ah mm [smacks lips] I don't know how do you say it in English
 = but ah like ah Inkeri [2.4] person [.2] people in Inkeri land
 and aahh
- S [mmm
- I = they were let to come in Finland in the nineties
- S right
- I and the government is providing them ah there ah here [breath intake]
- = accommodation and social security and [.6] so on
- S but but ca- can the- the [.3] ah [.5] Russians [.4] for example from Moscow
- = [breath intake] St Petersburg or ?{Budrek} [.2] can they come to Finland
- = quite freely
- I yes if they have roots [.8] in Finland
- S so th- they have to have roots
- S that's
- I [yes
- S = the main thing
- I [,] yes the foreign office is ah [1.0] aaahh is [.8] studying every case [.4]
- = every person before they can come
- S right [1.5] [clears throat] do they find that [.7] mm is there a big culture shock when
- = they come to Finland
- I yes enormous
- S yes

I I think aah well nowadays it's not that big ah [.6] than maybe ah three four years ago
 S yes
 I but aahh anyway [.9] th- the language is total different
 S mmhm
 I and usually they they don't speak any Finnish even though [.5] they [.3] they may
 = say they are Finnish
 S yes
 I they have had Finnish passports in [.4] in in Soviet Union or a Finnish nationality
 S do you think this was a good ah decision by [.9] Koivisto [.3] to allow these
 = people [1.2]
 I well for me [laugh]
 S [well [laugh]
 I = it's
 S [okay personally yeah there's [laugh]
 I [laugh] ah of course ah but the time was {diffent} then
 S [yes
 I there was open {vacancies} in in Finland
 S [uhmm
 I [,] a lot of jobs
 S yes
 I and so ah [.7]

I and [.8]
 S [uhm
 I = not such a huge unemployment as as right now
 S [so he didn't foresee the future very
 = well then
 I [yes it was I think it was a [.3] a mistake
 S yeow [.4] yeow [.5] what about these ah
 S [1.5] ah uh how could I explain
 S = / the nouveau riche
 S could I translate these as the [.2] Moscow [click] doing all their shopping
 = in Helsinki
 S mhm
 I has this had a great impact on [.7] Finland 'd you say
 I [.8] ahhh [.5] well I think uh [.3] well in in [1.0] uhhh [smacks lips] eastern parts
 = of Finland
 I I lived there for four months [.7] on winter [1.0] last winter
 I and uhm [.8] I think it's [.3] ah they are quite [1.3]
 I ah well [.9] you can see them very much ah [1.7] in the cities
 I and they are buying a lots of stuff
 I and there are lots of new shops [1.5] ahh for them [.7]
 S [uhmh

I = especially for them
 I and so they [.3] kind of provide the economic [.8]
 S [yeow
 I = situation there
 I and help people to get on
 S alright [.7] I [.9] just going on [glottal stop] on to this
 I [uhmhh
 S = it's [.3] it's sort of continuing from what we're talking about
 S ha- have you ever lived [.7] abroad
 I [breath intake] well [.5] ah I haven't li- ah lived there ah for a long time
 I but I have lived in Russia for [.7] few months
 S [yeow
 I = at a time
 S] so was this part of your studies or was it
 I [yes yes I had this uhm { obli'gatory } [.8] ah
 = studying in Russia
 S [uhm
 I = in ninete- in ninety-two
 S right [1.3] okay if you think of [1.8] obviously yo- you're applying for [.2] for the
 = Foreign [.5] Ministry [.6] Ministry of Foreign Affairs whatever they like to call it
 S and you might get [1.7] sent abroad
 S where would you like to go [.3] is it if not Russia [.7] let's say

I [1.8] well ahhh [1.2] I ah I've been to China once
 I and [chuckle]
 S [right
 I = eastern part of the world
 S [right
 I = interests me most
 S [yeow
 I but of course I'm ready to [.5] go ah travel anywhere and
 S [yeow
 I = live anywhere
 I because uhm I find [swallows] all the world very interesting [.9]
 S yeow [breath intake] let's assume that you're you've been asked to go abroad
 S what preparations would you make before going
 S [breath intake] you know personal preparations [.2]
 S what would you like to [breath intake] do before [2.0] the trip [?? weariness]
 I uhm well [.4] I think ahh it's important to know something about the country and the
 = people and the history you are going to
 I and uhm it's not [1.1] very bad to know so- some of the language too [.6] the basics
 I maybe [breath intake] basic two [.3] hundred words
 S [chuckle
 I something like that
 S [[c h u c k l e]

I [and of course very important to know all the manners [.3] and
I so that you you don't make all the horrible mistakes
S [uhm
I = at first [swallows] first day
S from your experience do you think that knowing the dos and don'ts of a culture [.9]
= can actually help you or hinder you [.8] when working in that culture
I [1.1] you mean ahh how in what way {hinder}
S ahh [1.4] they can be [.7] by knowing them you can be put a-at a disadvantage
I [2.3] ah well I think it helps you [rolls eyes] [.7] if you know [.2] know
=cultural differences
I although maybe you can't [1.1] ah can't follow them [.8] in every occasion
I you have to be ah neutral [1.2]
S right
I ah you can't take a side in every situation [?? any]
S right
I but [.6] I can't see why it why could it be bad [.2] if you know about things
S mhmm [.7] no just ah there's just one interesting article where they [.5] said that by
= knowing the dos and don'ts of a culture you feel [breath intake] people going
= there feeling that [breath intake] they know all about the culture
I uhuh
S so therefore [1.5] they tend to rely on that limited knowledge
I uhuh

- S and just miss everything else
- I mhm
- S what would be your impressions of that
- I well [.5] I I think [exhales] there are people [1.5] different kinds of people
- I and ah [breath intake] some of some of us ha- have ah different kind of [.5] attitude
- I and ah you can't change the person's attitude [.5] that much
- S mhm
- I if you are uhm if you are open-minded and you [.3] you want to learn about things
= you'll soon find out [.6] ah it wasn't right what you read [.4] in a book or [.4] heard
- S [1.2] okay
- I [1.1] so I think ah some knowledge is always [.7] anyway better
- S [mmm
- I = than no knowledge
- S yeow [.3] what about the people that you work with [.3] these Russians
- S have they [breath intake] adapted to life in Finland would you say rather well
- I [1.5] no unfortunately I I wouldn't say that because uhm [breath intake] there are
= many people who have lived here for [.8] many years [.5] already for five years [.2]
- I and they still need me as an interpreter when they go to social [.4] social
= administration or ah at schools or ah [1.6]
- S [mhm
- I so uhm I think we have a very big this huge gap between Russian and Finnish cultures
- I I can't always understand why it is so big

S [yes

I [.9] because we're that much different from each other

S where would you put the blame if any blame co- [glottal stop] could be put

S is it Finns not accepting [breath intake] or the Russians not accepting

I I think we are both to blame

I but uhhh [smacks lips] part of it ah part of the problem is probably because

= uhm the Russian were ahh this great nation

S mmmm

I they used to be this great nation

I and uhm I don't think I don't think ah that most of them quite have realized that

= they have lost the position [.9] they had

S yes

I and ahh so they can't go everywhere

I and ah not all the people know Russian anymore [.8] like it was in the Soviet Union

S so any- anyway th- be- from what I know they are a very proud nation

I yeow

S and now they've lost a lot of that

I yeow

S [that pride

I yeow yes I I don't know what's the problem because ahh [breath intake] ahh many of

= them are are very well educated

I and [.7] they know their profession

I they could work here easily
 I but [.4] Finnish language is [.7] the problem
 S yeow yeow
 I they can't solve [chuckle]
 S [1.4] [exclamatory] interesting thank you
 S uh uh I think that's it
 S time -s has gone [.4] so quickly
 S there was this really
 I [chuckle]
 S do you have any questions any uh
 I uh no not actually [laugh]
 S alright I think that was it
 S so good luck with the rest of the [2.0]
 I thank you
 S I'll just turn this off
 S [1.2] the rest of the [.7] test
 I right thank you

 total time of interview 11 min. 9 sec.

INTERVIEW 2

dyad identified as: Kaija Karkkinen-Smith (I) and Jack Doe (S)

Impressionistic notes on the videotape:

[sound quality on the recording is not very good, perhaps a microphone on the camera, so, e.g. question intonation not noted]
[I is very competent, a researcher, University of Helsinki, visiting scholar at the University of Pennsylvania, who lives most of the year in the US, married to an American; S appears a bit nervous, as seen on screen, twisting a pencil, etc., perhaps daunted by her]

S you know this lady is doing the [.2] the [gesturing at camerawoman]
 I [the tape
 S yeow okay [breath intake] and [.4] I also have to ask you to introduce yourself
 = just [.7] ah especially the name [.8]
 I [mmmm
 S = and [.2] then we'll have a chance to [implied: do the interview]
 I [mmmm
 S [[sub] ??
 I [mmmm
 = I my name is Kajja Karkkinen-Smith
 I and I'm working as a research at the University of Helsinki [breath intake]

S mhm
 I and live here at Helsinki presently
 I but ah [.5] most of the year I am living in a- America [.2] near Baltimore
 S uhuh [.2] [?] but you are still attached to Helsinki
 I yes, I am I ha- I am working for for fi- Finnish university but in in America
 = doing a research project there
 S [how does that work out
 I it works it works out well
 I in the sense that th- research you can carry with you
 I you know if you have big luggages [chuckle]
 I you can take your big luggages and carry
 S [yeow
 I = them [breath intake]
 I it works well in that sense
 I and I have computers
 I you can also carry those
 S yeow
 I and [.6] of course communication [.6] over [.5] across the borders [.7] is
 = always [.3] fruitful for a research project
 S yeah
 I but mainly on on the other hand you are always {away} [breath intake] from
 = your main project

S yeah
 I so so then [.2] in a sense there is positive and negative sides
 S okay yeow [.5] are you involved forrr professional reasons or personal-
 I [in Baltimore
 = I'm for [.5] ah for reasons of because I'm married to American
 I so ah he he lives in Baltimore
 I but I'm working at in ah Philadelphia at the University of Pennsylvania
 I and we are very close [breath intake]
 I so I can commute
 S okay okay [.4] how how about connections with the University of Helsinki
 S is there any [.5] connection on a [.3] higher level
 S or is it [.3] just your own [.2] arrangement
 I [it's it's that connection is mine in a sense that
 = it is-
 I [my research project has agreed to finance me [breath intake]
 S yeah
 I and they are very positive about me going abroad and
 S [yeah
 = working there
 I but [.4] ah I have myself found a place [.2]
 I I'm a visiting scholar
 I so it was you know [.7] it was a question of my own initiation

= in a sense [breath intake]
 S [yeah
 I = to go and find a place to visit and to work there
 S dare I ask you what is your area
 S I was looking [gestures to papers]
 I [it's medieval history
 S [medieval hist-
 I [= history of medieval ah ah religion +z
 = gions and especially that of lay laywomen
 S mhm
 I so I'm studying the religious ah life of a medieval lady
 S and is Baltimore i- in some sense a center or or ah I was thinking
 I [Baltimore would
 = be a cr-
 S [sub] ??
 I [yeah yes it they have a John Hopkins University which has ah several
 = very distinguished professors [breath intake]
 S mhm
 I = and for in that sense it is a center of its own [breath intake] because of its professors
 = and because of a very distinguished publishing company
 S okay
 I there again on the other hand for myself there isn't [.2] uwpf

I Johns Hopkins i- is focused on a different research field
 I so it is because of my husband that I am there
 S æh okay [breath intake] and [.4] maybe that [1.1] well that that you are living in
 = Baltimore and your [.4] other choice of topic are somehow related to your
 = professional area
 I [yes
 S = in that sense
 I [yeow
 S = and so shall we move on to th- the second of your topics
 I yes
 S and I'd like to look at this one a- around th- the theme literacy
 I mhmm
 S and [1.2] I'm going to start at the bottom or- at least in terms of the lists here
 S and ah ask you can you make a definition of literacy [.2] because I think
 = actually what your opinions are [.6]
 I [mhmm
 S = of literacy ah depend very much on what you understand of it
 I mhmm and especially in modern times it is very interesting issue because
 = the literacy has changed from the novels [breath intake] of ?poesy towards ah
 = poe- poetry towards something which is more to do with also with computers
 = and the computer literacy [breath intake]
 S [mhmm

- I = and how does the traditional [.5] old you could maybe say old-fashioned
 = literacy how does it react [.3] to the modern way of reading [breath intake]
 = which is nowadays more- it's ah it's focuses on a shorter reading sessions in a sense
 I you read something short from a computer or from TV or from a newspaper
 I [breath intake] and it's very interesting to see [.5] how this traditional [.5] reading and
 = ah longer narra- na- narrative cycles reacts
 S [yeow
 I = to this kind of different way of reading
 S okay
 I so I would say that literacy you have to define it from the two [.3] angles
 I the more ah traditional ah [.2] and a more modern one
 S mhm
 I [,] and what should be included
 I a lot of ah ah learned people who [1.3] [sub] from among /who are who are ah more
 = for old-fashioned reading for reading novels [.5] of course feel very [chuckle]
 = upset about the modern way of reading [breath intake]
 I I do too
 I but then again I think you have t- to have both [.3] in
 S [hmpf
 I = a definition
 S [okay I mean [.3] would you accept that literacy can be [?] measured [1.6]

- I it can be measured
S [mm
I = in its very basic forms
I it can be measured if you have a literacy rate for example of a certain [breath intake]
= po- ah population
I and you can see ah if it is a literate or non [breath intake]
S mhmm
I but aahhh and I think that you can you can ah [.3] measure a basics of it
I yeow
I and then again the deeper understanding [breath intake] I think is ah [.4]
I that would be awfully hard as anything quantitative [breath intake] is very hard
= to me-
S [ye- ye-
I = really measure
S [yeow [.2] I was thinking th- that it's maybe easier to measure
= [sub] ?? /one of the old-fashioned definitions of it cause see [.3] what's
= the difference the more modern the more
I [yes and if you would want to measure ah
= literacy as a ce- certain critical eye [breath intake] as ah something that when
= you are literate you also need to be [.6] ah critical towards what is ah in the text
I that is very difficult to measure because that's more kind of like a cultural literacy

I how are you able to understand the message [breath intake] the symbols the
= hidden meanings [breath intake]

S [mhm]

I = how to measure this

I but of course it is important because that's an important part of our culture is this
= kind of hidden [1.4]

S [yeah

I = levels

S yeah [.8] are these themes or topics that come out from your own research or [.6]
= in a sense that [.2] ah presumably writing in medieval times [.2] there was the
= same [.9] an audience [.2] an author [.2] and to some extent you know
[yeah then it's

I = [.2] ah in a certain sense yes

I and for the medieval ah literacy it is a more a question of reading and reading
= aloud [breath intake]

S [mhm]

I and then that is again a different issue how how

I but in medieval literacy it would be something to do with ah public gathering [.8]
= someone reading

I and then the understanding would be more something about [breath intake] whole
= understand what is read to you [.5]

S [yeow

- I and even the researchers ah point out that it was only in the fifteenth century [.3] that
= people ah learned to read quietly [breath intake]
- S [yeah
- I = before that the reading process went through that even the one who knew how
= to read they needed to read it ah aloud
- S yes yes didn't know that
- I yeow
- S how about one of the other points mentioned here that illiteracy sorry literacy
= is something that develops through life
- I mhmm
- S you have an opinion
- S [are you agreeing with that
- I yes I I think it's also a question of your be- your vocabulary [.3] your aahh again this
= sense of how how critical you are in your understanding of language
- S mhmm
- I and ah [.6] and unfortunately it tends to be that a lot of people who in the [??] are
= able to profit from a very rich linguistic surroundings
- I are able then to go to ah ah you know fo- [.7] are able to develop a writing
= skills [breath intake]
- I their discussion skills far better than those who [breath intake] have later start
- S [yes
- I [,] so I think it's go- it's something you get from your mother's milk

- = not not completely [chuckle] but symbolically speaking
S [staring at papers] how about then the the relationship between literacy and
= [.2] success
- I [mm
S = is that automatic
I it is it is very [.3] the relationship -ip is very strong yes [breath intake]
S [okay
I = and I think especially in -n nowadays when the image [.6] is such an important part
I = it can be the ah extr- image of your- [.3] exterior image your looks and so
S [mhm
I = but ah also how you know how to use the languages
S [mhm
I = and again foreign languages [slight chuckle]
S [mhm mhm [.6] do you think there's a significant
= difference between Finland and America in this context
I in a in a sense that Americans aaaarre ah better in making aahh very [.8]
= ah [breath intake]
I [parenthetical] would you say
I making ah statements ah in a very [1.1] in a form that is very very selling
I [very ah very æh [.3] very [.2] somehow catches the idea very quickly
I I think Finns ah still need to learn that technique
I something which goes to the market

- I something which helps [breath intake]
 S [yes yes
- I = [,] and that can be used in a very spiritual meanings [1.2] not only in the
 = commerical world
- S yeah [.7] so [.3] is that a form of literacy or is it just a skill
- I I think it's it- an American skill especially American skill
 S [right [1.1] -
- S so how about literacy in in America as a [.6] [sub] you know /whole
 = community the whole country
- S is it- nn is there any comparison with Finnish [.2] literacy
- I it's a it's an interesting question because I I think in America [breath intake] in
 = America you will reaa- or everyone knows that the the education is so
 = two ah [.9] is ah two-fold
- I you either have a public or private education [breath intake]
 S [mhm [mhm
- I = and this is I think the the social issue that is very difficult to solve
- I [because now that I meet people it's quite [breath intake] you quite easily
 = realize what kind of basic education they have had
- S [mhm
- I = [and unfortunately it is a case that [breath intake] the private education is so
 = much better -an as the public
- S [mhm

- I and here in Finland luckily the ah the public education is rather standardized and
= quite high
- S yeow
- I and I think th- in America the social issue is between [.4] how to [1.0] concordate the
= public and private
- S so [.2] so can one actually compare literacy in the two countries
- S do you think that [.6] they have the same type of literacy [.4]
- S or is it somehow [1.3]
- S shouldn't we really bother to compare
- I I I think one may [.3] one can compare it
- I but they are very different cu- cultures in their way of using [.7]
- I and how they react to the language [breath intake]
[mmm
- S = and ahh but I would say it's still com'parable
- S [sub] yeah yeah [.6] /so presumably Finland comes out on top in terms of
= total literacy
- I [yes that is
- S =overall literacy in a sense ah a democratic learn- learnedness [breath intake]
[mhmm
- I learning democratically
- S [mhmm

I = and the basic educati- and pa- libraries
 S [mhmm]
 I = is a very important issue here
 S [mhmm]
 I = the Finnish and also Scandinavian public libraries [breath intake]
 S [yeow]
 I = are phenomenal
 I then Baltimore is a city that reads
 I they advertize it [chuckle] already
 S [yeah]
 I = /the city that reads
 I and you see it in a pages [breath intake]
 S [yeow]
 I = of the public bus stops
 I there is a- the city that reads er- ah always in the benches
 I and Baltimore actually has [.6] very good public library system
 I but that's rare [1.5]
 S [yeah]
 I = [sub] in all America [1.8]
 S alright [.8] this has been an interesting discussion [.5]
 S aaahhh [.4] I don't think we need to cover the other topic

S [sub] I mean / that was a possibility

S but I think that we we can

[tape cuts off at this point]

total time 9 min. 59 sec.

INTERVIEW 3

dyad identified as: Liisa Lintunen (I) and Harry Doe (S)

Impressionistic notes on the videotape:

[apparently there is some problem with the video camera, or something else, which prompts the interviewee's first statement]

- I everything's falling apart
S [breath intake] yeah [.7] the whole world [deadpan response]
I [subvoc] yuh
S right okay so I'm Harry Doe [?] you are
I Liisa Lintunen
S [?] Lin- Lintunen
I [Lintunen
S [Li-
S [um okay
I [okay
S yeah I have I have to work on these names a little bit
I [yes
S = sometimes [1.0] [breath intake]
S] okay right so [?] why are you here today

S [ʔ] why are you doing this test
 I aa-a first of all I like English language
 I but I'm also
 S [uhuh
 I = applying for the Foreign Affa- Foreign Ministry course
 S [uhuh
 I = to be employed there so
 S [[ʔ]Foreign Ministry course
 I yeah yes
 S [ʔ] which is
 I [,] which is [.3] can't remember what 'tis in English
 S [sub] can't remember [.3] well no matter
 I [[breath intake] mmm
 S okay but some kind of course
 I yeah [.7]
 S [yeah
 I = to be employed [.6] abroad
 S [alright
 I [,] later on [.5] after working
 S [alright
 I = one and a half years or two years in Finland
 S [alright

I = in the Ministry
 S] alright we're talking about some kind of diplomatic work then
 I right sort of [chuckle]
 S [uhuh so going where do you think
 I I hope to go
 I I have lived four years in Tanzania
 S [?] Tanzania
 I yes we came back in March
 I so [breath intake] I've always been interested in going abroad and working there
 S mmhmm
 I but uh we'll see
 I because it's quite hard to get [.3] to this [.5] course [.6]
 S [mm
 I = or even [sub .6] employed for the [1.3]
 S [mmhm
 I = Foreign Ministry [.2] in Finland
 = because people are so hi- highly educated [.6]
 S [mmm
 I = have a lot of experience [sub] like from languages and cultures
 S [yeow
 I = nowadays

S alright [.7] well [.2] well we'll be
 I [yeah
 S = optimists
 I [,] yes we can be always [chuckle]
 S [okay so [sub] I mean [.2]
 = you've lived in Tanzania for four yearssss
 I [yes
 S = so [sub] both of you [.2] know about living in a foreign culture
 I ah a little bit you it's know only a scrap [.3]
 S [?] a scrap [?] four years [?] only a scrap
 I [I think because ther- ah it's very different
 S uhuh
 I [,] aaahh we've al- also been eleven months in Birmingham
 I and
 S [uhummm
 I = I was [.3] [sub] believe [.2] six weeks in Vienna [breathe intake]
 S [uhummm
 I = and three months in London
 S mhm
 I but I still figure that in order to learn to know the people well
 S [mhm

I = and the culture you have to spend years [breath intake]
 I and I say for us it was a scrap because we lived in an African village where
 S [mhm
 I = there has been missionaries for [.5] one hundred years
 S mhm
 I and they were Germans and Finns and Swedes living in same village
 S [nod]
 I and we sort of lived in our own [.8] community
 S oh right
 I and it's very much in Tanzania that the Europeans they live
 = in their community [breath intake]
 I and the people from India and Asia living in their community [breath intake]
 I and then the Africans
 I and ah the sort of [.3] cooperation and communication between these [.7]
 S [alright
 I = it's not that easy
 S so yo- you would say you were not integrated into the culture
 I we were partly [.4] and time to time
 S mhm
 I but still you belonged t- in- to many groups
 I or you have a lot
 S [right

I = of expectations
 I and you can't live as the Africans lived [.9]
 S [mhm
 I = because we can't live in a clay hut with three children [breath intake]
 I and we wouldn't live there even if we didn't have [1.0] the children
 S mhm
 I [,] you know we had the cars
 I and we had more money to spend [1.9] [breath intake]
 S [mhm
 I so [.3] we don't integrate [.4] totally [.8] to the culture
 S it sounds like you don't integrate at all almost
 I [breath intake] in a way you could say that {thet}
 I I mean I sometimes I I felt that we lived like the colonialists [1.4]
 S [mmm
 I = years ago uh [1.3]
 I but I mean we did integrate because [sub] when I mean
 I = when I was teaching in the school
 I [aside] my husband was a science teacher in the secondary school
 I unfortunately we spoke a lot of English because we had to teach in English
 S mhm
 I and I integrated int- the culture when I was traveling with the bus [breathe intake]

I and
 S [mm
 I = going my own way [.5] with the localll [.6] you know the [sub] well
 = local transport
 S [mhmh
 I = for example
 I and we went we went to the weddings [.7]
 S [mhmh
 I = or to the feasts of the school
 I it was you know popping in and out of these cultures
 S [sub] well [.2] still sounds to me like you were paddling in these cultures
 I yeah we yes [.4] that's true we're going out and in
 S mhmh
 I and living in different worlds at the same time
 S mm
 I and I think it was quite hard
 I [breath intake] but it was interesting
 S mhmh
 I and ah [.2] I think that [.4] we've got a lot to learn from them
 S] right [.2] but [.3] wouldn't you say

S that there was a world of difference between say [.2] me who comes to Finland [.9]
 I [mhm
 S and [.9] unless people speak to me I could pass in the street for a Finn [breath intake]
 S and you going to Tanzania
 S and [1.0] being an outsider who's come temporarily
 S very clearly not a local
 I [mmmmm [1.1]
 = th- there's a big difference [.9]
 I did you ask that
 S [ah that's right
 I = yes [.5] [breath intake}
 I because first of all I was always watched [.8] and stared at [1.4] [breath intake]
 S [mm
 I = but I was also helped
 S [mhm
 I = but I I was always a foreigner because of my
 S [mhm
 I = color
 S [mhm

I = even though I know that a lot of Finn- Finnish p- people
= who speak Swahili perfectly
S [mhm

I = and they say
I [interjection] they're very polite
I and they say that you're like one of us [.3]
S [mhm

I = a local person
I but then in the end it comes [.6]
I th- the
S [m
I = reality is
I = that the people don't get that much friends with the local people
S [mhm

I = I heard one lady who has lived there ten years
I that she has one friend [1.5] [breath intake]
S [mhm

I = and [.2] there is [.6] there are many gaps ge- to cross over
S] is it a question of money as well
I [aw that's the main question

I = th- the biggest problem perhaps
 S [uhuh
 I = that if we lived [1.8] as they lived there might be [.5] some kind of a way [.8] to go
 I but I figure that [.5] sometimes the color [.6] fazes out
 I it disappears
 S [mhm
 I = we didn't see that they are Blacks
 I and we were Whites
 S [mhmm
 I = and momentarily you could [1.1] be brothers and sisters
 I when- [.2] and the cases when it was [.7] more [.7] [lip implosion]
 = most [.4] ah visible [1.0]
 S [mhmm
 I = or [.8] more easily experienced was when I was needing help [breath intake]
 S [mhmm
 I = when I was asking help or [.5]
 I or [.7] when th- when changed their roles
 I I wasn't not the- I wasn't anymore the rich European with the car
 S [mmm
 I I wasn't hitchhiking
 I but if I was using a bus

S yeah
 I we had the same [.5] you know the same thing in front of us [1.3]
 S [mhm
 I = and ah [.5] there was one Muslim man he asked me money [1.0]
 S [mm
 I if I could give him
 I and I said I don't have
 I why don't you give me [.7]
 S [mhm
 I = and he was so surprised that he st- that he burst out into a laugh
 I and says [breath intake] everyone around him
 I that look this lady's asking me for money
 S [mm
 I = and was so happy
 I that i- i- it's very [breath intake] you know instant cases [.5]
 I and then we feel that we are both human beings [.4] equal
 S [right
 I = qu- quite the same
 S [mhm
 S you mean there might be an [.3] element
 S that sometimes [.5] the Africans feel [cue for c-unit change from videotape]
 S that you are being patronizing

I [breath intake] [forcibly exhales] yes and I I think in Tanzania especially I was hurt
 I compared to Ethiopia and other countries
 I that we felt [.6] that there's [.4] in background
 I or deep
 S [mhm
 = inside in feelings
 I there's a- a ?still ?member even though we we have not been colonializing [.5]
 I [mhm
 S = Tanzania
 I but we realize in Europeans there is a racism towards us
 I [[?] mhm
 S but is there an element of status
 S you've come [.2] to give [.8]
 I [yes
 I = and you've come to teach [.2]
 S [yes we've [.3] come to give what
 I [yeah
 S = you know we [.3] think we're a little bit better [.5] [breath intake]
 I [mhm
 S = or we have something to give to them
 I [mhm
 S

I = but I think that if we have an attitude that they have a lot to give to us [.6]
 S [mmm
 I = at least in mission work it's [1.2]
 S [mhmm
 I = gradually changing [breath intake]
 I that people are coming here and preaching the gospel here
 S [mhmm
 I = and pe- ar- the African [.6] priests
 I [breath intake] ah [.7] so we think more that we can share
 I there's a lot to learn from the community and society life [.5] and social [.7]
 I what we have now here [1.1] established
 S [mhmm
 I = new things you know
 I you will have to learn work gr-
 S [mhmm
 I = working in a group [.5] and take care of each other [.9] and pay attention to your
 = neighbor and so on
 I these things [.6] are social matters
 I I think we w- would have a lot to learn from them
 S mhmm so I mean does [.4] traveling and living abroad [.7] actually
 = help [.3] global understanding

- S or is it illusory
- I I think it helps [.3] a lot because when I'm seeing someone [.7]
[mhm
- S = where he lives [breath intake]
- I you can understand better
- I and I feel it's good for the Africans come here
[mhm
- S = because then they understand where we come from [breath intake]
[mhm
- I you know what's surrounding
- I and what we have in the shops and the material and [.6] and [.3] the technical things
= and the communication facilities and so on
[breath intake] I think it does
- I but there's always a danger th- that it [.6] makes us more [.5] {prejudiced}
[mhm
- S = and strongly [1.2] ah to op- [.8] to- not to oppose [.7] but to postpone our own ideas
[?] postpone ba-
[not postpone it's to [.5] to emphasize and believe
[right
- I = in our own ideas more [1.2]
- I if we still- we don't- we see them- the others very differently and think
[mhm
- S

I = they are- they can't change [breath intake]
 S [mhm
 I = and we think all the Blacks are [.3] you know lazy men [1.2]
 S [mhm
 I = for example
 I so I mean [.3] can it sometimes happen [.8]
 I ah I've spoken to people who've for example for periods- [.7] ah that actually when
 = you go abroad your prejudices are sometimes confirmed and made worse
 I yes that's true [.7]
 I it can
 I it depends I think it
 S [[?] depends on what
 I it depends on [.4] yourself [chuckle]
 S [uhuh
 I = ri- I- I think to a great extent [breath intake]
 S [mhm
 I = it depends on [.3] perhaps on on your background [1.6] in the child-
 S [mhm
 I = and the feeling that [sub] when the-
 = yourself you're 'content with yourself
 = [breath intake] you know when you have problems th- [.4]
 = with your [.2] personal life in your personal life

S = you may reflect [.4] them to these people who are different
[mmmm

I = who behave differently

I [breath intake] so [.8] I think it's a matter on a personal level

S mmmm

I = ahm [.8] I had an idea I forgot [.2] what was it [2.3]

I ah just continue [.2] i- if it comes up I'll tell [giggle]
[okay wh- what about this concept

S people sometimes talk of a global village

S we're all on the same planet [.8]

S I mean are these meaningless phrases or [1.3]

I ah [.7] no I don't think they are quite meaningless

I we were in a concert [.8] the day the day before [.7] yesterday

I [breath intake] and I feel that especially in Finland people don't want to go [.3] out
and they want to be where they are other people [breath intake]

S [mmmm

I = so just the experience that we belong to a bigger group than my family
that our school [.9]

I [mmmm

S = for example we have these festivals [.9] when it's warm and we can go out
people are very

S [mmmm

I = keen on going [.2] nowadays to these
 I and also that we have the TV
 I and we have a lot
 S [mhm
 I = of big happenings with millions and thousands of people [breath intake]
 I th- there's a need that we [.4] want to belong to the earth [.7] to the
 = global [1.1] ah community
 I [1.2] ah [1.3] I mean- I know- I I'm thinking as a matter of feelings did [.4]
 I you know we need to feel that
 I [breath intake] and we fe-
 S [mhm
 I = experience the feeling
 S [mm
 I when we goes to these mass [.6] happenings [breath intake]
 S mhm
 I = like concerts [1.1] or ah [.6] festivals somewhere [1.7]
 S yea but surely when most people go abroad they want to see something [.2] different
 S let me give you an example
 I [mhmm
 S = if you were to go to London to see British soldiers in
 I [mhm

S = big red tunics with big high hats
 I [mhm
 S = marching [1.1] [breath intake]
 I [mm
 S = ninety-nine percent of the British soldiers aren't like that
 I right, yeah
 S [but they go to see the difference [1.4]
 I [yes
 S = not the normal ordinary people
 I eah [2] I think that's ah {phenomenon} which is it's actually
 I we have the TV [.7] and the books [.9] which add to tourism [.8]
 S [mhm
 I tourism [.2] where there are people who want to look as an outsiders
 S mhm
 I then there are people who want to live and go and live there
 S mhm
 I [breath intake] and I think they have a different view [2.1]
 S [mhm
 I = to live with the English people
 I there are [.6] there are few them left
 I so if of course it's an experience

I I mean the TV is not giving us enough
 S [mmmm
 I = it used to give
 I now we have to go and [breath intake] and see where it's happening
 S right
 I and still being outsiders [.4] because it's easier [1.1]
 S [mm to be an out-
 I [to be an outsider
 S [alright
 I = you you can enjoy [sub] ?? you don't have to take part
 S mmmm
 I and your- [.4] you can go through your feelings differently [1.1]
 S you don't get hurt
 I no [.8] and you don't get shot [.5]
 S you don-
 I but you can see a crime [chuckle]
 S [yes that's true [1.1] oh well [.2]
 S that's very interesting
 I yeah
 S = to hear about all that
 S thanks for coming in today [shaking hands at this point]
 I [thank you

S and good luck with the rest of the test
I [okay [breath intake] thank you very much
S thank you

[I: off-screen: bye]

total time 12 min. 50 sec.

APPENDIX 2

NATIONAL CERTIFICATES: SCALE OF ORAL
PROFICIENCY

THE NATIONAL CERTIFICATES: ASSESSMENT CRITERIA FOR THE ORAL TEST - ADVANCED LEVEL

VERSION 17/3/95

Overall	Pronunciation and prosodic features	Accuracy of structures and vocabulary, and idiomacy	Fluency (narrow definition)	Discourse skills, Linguistic appropriacy	Presentation (tape-mediated)
8	Approaches native speaker competence. Speaks very fluently and is able to use idiomatic expressions. Only occasional non-native features, e.g. has to search for an appropriate expression or has a very slight foreign accent.	Phonology is accurate. Liaison of words is smooth, demonstrating broad skill in application. Tone group and word stress, intonation, and the use of other suprasegmental features regularly contribute to the effectiveness of the communication.	Extensive and accurate use of vocabulary in all tasks. Very often demonstrates exceptionally good use of vocabulary. Is able to convey nuances and use idioms with ease. Complex or rare structures may cause slight problems in conceptually demanding situations. Good command of structures. Use of language is very idiomatic on the whole.	Dysfluency due to missing language ability cannot really be detected. The examinee may search for 1-2 expressions, but handles this smoothly and effortlessly. Hesitations and pauses are natural and result from having to deliberate on subject matter. Automaticity/pausing natural and appropriate.	Presentation is well-organized, clear, and advances logically. Main points are given sufficient emphasis. Structures speech in effective and versatile modes.
7	Handles public speaking/presentation situations well. Speaks clearly and fluently. Some features, such as intonation and stress, may need improvement. Participates actively in conversation. Able to use various language registers appropriately. Able to convey nuances fairly well.	Phonology is to a high degree accurate. Liaison of words is for the greater part smooth. Tone group and word stress, intonation, and the use of other suprasegmental features does not at all times consistently contribute to the efficiency of the communication.	Sufficient and accurate use of vocabulary in all tasks. Exceptionally good use of vocabulary at times. Conveys nuances and uses idioms and culture-bound expressions with relative ease. Good command of most structures. Grammatical/syntactical errors are few and occur primarily in situations demanding language use.	Able to adjust language rather well according to situational needs, and usually able to choose the suitable degree of formality and politeness. Can begin and end the conversation in a smooth manner. Takes turns, reacts, changes the subject, and covers various topics mostly in both a flexible and a natural manner.	For the most part clear and well-organized presentation. Structures speech in rather effective and versatile modes.
6	Copes well in everyday occupational and other speaking situations and is able to take initiative. Can make slight modifications to standard expressions according to situational needs, e.g. mitigating a request or a demand. Speaks clearly and accurately, making comprehension effortless. Some features, such as intonation and stress, may need some improvement.	Phonology is accurate for the most part. Is able to use intonation and other suprasegmental features rather appropriately (e.g. producing a distinction between a question and a statement, or a request and a demand), but on the whole the production of suprasegmental features is not consistent. Word stress mostly on the correct syllable, with the exception of less common vocabulary.	Vocabulary used on specific tasks satisfies task requirements, but demonstrates neither obvious limitations nor special merit. Some inadequate use of vocabulary. Has a good overall command of, but makes errors in more demanding structures. N.B. At Levels 6-8, errors (mostly in vocabulary) only seldom create comprehension problems.	Relatively effortless, smooth flow of speech. Search for appropriate expressions causes slight unnatural pauses or hesitation at times.	Fairly clear presentation. Structures speech for the most part in effective modes; slight repetitiveness of cohesive devices.
5	Copes well in everyday speaking situations and fairly well in unprepared-for communication situations. Utterances are fluent, connected well and of appropriate length. Able to present and supply grounds for views and opinions. Can make some distinction between formal and informal registers. Attempts to produce variety in expressions utilized may sometimes lead to incorrect forms and impede comprehension.	Individual sounds are usually accurate, and word stress mostly on the correct syllable. Tone group stress, intonation and/or rhythm of speech recurrently deviates from the norm of the target language and may be in conflict with the informational content of the utterance.	Mastery of vocabulary may not be quite sufficient for conversation on the given tasks. Occasionally produces some accurate and fairly detailed expressions. Commands basic structures. Is able to correct simple errors, but cannot correct complex ones.	Makes an effort to adjust language according to situational needs, but outcome is fairly often inadequate. Usually able to differentiate between formal and informal registers. Manages fairly well in conversation, though demonstrates awkward usage rather often (e.g. reactions may be atypical).	Only fairly clear presentation. Some sections may be structurally unclear, or some main point may be insufficiently emphasized. Structures speech fairly appropriately, but use of cohesion markers is e.g. repetitive.

APPENDIX 3

NATIONAL CERTIFICATES: INTERVIEWER TRAINING GUIDE

NATIONAL CERTIFICATE
LEVEL OF USAGE INTERVIEW
NATIONAL PROFICIENCY LEVEL
CERTIFICATION/FINLAND
SECOND LANGUAGE/ENGLISH
INTERVIEWER TRAINING

Training Session
Feb. 23-24, 1996
National Board of Education

CONTENTS:

INTRODUCTION
THEORETICAL BACKGROUND
ASSIGNMENT OF ANALYTICAL RATINGS FOR
SPEAKING
LUSI WALK-THROUGH
ROLE OF THE INTERVIEWER
APPENDICES

INTRODUCTION

This test of speaking (2 subtests of the test battery, each comprising two tasks) aims at determining an adult examinee's oral language proficiency. The face-to-face subtest consists of two tasks, interview and paired discussion, and there are two speaking tasks on the tape-mediated subtest. However, the overall mark on the test of speaking is weighted in favor of the face-to-face test. Therefore, it is extremely important that the interviewer is able to elicit a sample of speech that fairly represents the interviewee's skill in the language tested as well as demonstrates where possible limitations in skill lie.

THEORETICAL BACKGROUND

LUSI (Language Usage Interview) is a proficiency test, as is the entire Certificate test battery. It is the aim of LUSI to determine the examinee's skill level at the time of the testing, irrespective of when, where, how, or under what conditions the language skills have been acquired. The examination measures functional general-purpose language proficiency. Therefore, the test is founded on theoretical constructs of what speaking skills are and what determines various skill levels in speaking. At the Advanced Level these constructs subsume: broad, flexible vocabulary; active, extensive, workable grammatical knowledge; accurate, clear pronunciation; effortless flow of speech; adept usage of discourse management devices; and broad knowledge of the appropriate use of language.

A high performance level in speaking exhibits a natural flow of speech, including natural hesitation phenomena. The interviewee uses a wide variety of individual vocabulary items, collocations, idioms and phrases, with no indication of difficulty incurred in even the most complicated grammatical structures. On the other hand, the interviewee demonstrates sensitivity to less formal oral situations which require the use of simpler structures and a more colloquial lexis. The spoken performance can be followed without undue effort, as the individual sounds, stress patterns and intonation all contribute naturally to the comprehensibility of what the speaker is communicating. In summary, the interviewee's language skills in no way hinder expression, i.e. on varied topics, more or less demanding, with a higher or lower degree of abstraction.

LUSI is criterion-referenced (CR), in contrast to tests which are norm-referenced (NR). This means that the performances are compared to the skill level criteria, not to each other. It is important to reiterate that the level of

ability constitutes the criterion, not the setting of a cut-off score for making decisions. There are no pre-determined proportions of examinees prescribed to reach any particular skill level.

For our purposes, the appropriate setting of the criteria is observable in the Assessment Criteria for the Advanced Level—Oral Test. Other applicable criteria are to be found in the Skill Level Descriptions and the Advanced Level Proficiency Descriptions. The Assessment Guide/Advanced Level: Speaking also contains a helpful section entitled "Comments on the Proficiency Levels", designed to help the interviewer keep in mind the overall structure of the exam.

ASSIGNMENT OF ANALYTICAL RATINGS FOR SPEAKING

The assessment of speaking proficiency on LUSI is arrived at by combining the 4 applicable criteria (Pronunciation and Prosodic Features; Accuracy of Structures and Vocabulary, and Idiomacy; Fluency; Discourse Skills, Linguistic Appropriacy) into an Overall Description of Oral Skills. In the Speaking Subtest, these are then combined with the assessment of the tape-mediated tasks (5 criteria, the above and Presentation) into a Skill Level for speaking.

On LUSI the assessment is a mixture of an analytic procedure where no overall grade is given and a holistic measure where no ratings for individual criteria are awarded. The interviewer's immediate assessment of the interviewee's performance is, thus, the most "holistic" of the speaking assessments. These assessments, 4 criteria categories, are assigned numerically. However, the most important numerical indicator the interviewer gives is the Overall Speaking Level. Further, it is very important that the interviewer records impressions of the interviewee's oral skills in writing. Afterwards, LUSI performance from each candidate is assessed by two independent video assessors.

Results are then compiled. The reliability of the mark assigned is ensured through: 1) the use of the Assessment Criteria for the Advanced Level-Oral Test, and 2) uniform and continuous training as well as the multiple marking of subtests.

LUSI WALK-THROUGH

LUSI has a conversational format, a 12-15 minute interview (interviewer with interviewee A then B) followed by a paired discussion (interviewee A with interviewee B). For the interviewer a framework for each examination is provided in the form of a protocol. The interviewees are divided into pairs in advance by each test center. Before the interview, each pair meets outside the interview room and decides, from a selection of three, which topic they would like to take up in the paired discussion (1-2 min.). After that, A is interviewed (12-15 min.) while B prepares for the paired discussion. Then they exchange places. At the end of B's interview, A is invited in for the paired discussion.

The interview consists of a warm-up, a thematic interview conducted by the interviewer (using the interview protocol), and a wind-down. During the warm-up the interviewer gains first impressions of skill level. The initial level of difficulty should be low enough for the interviewee to gain confidence in the interview situation. In addition, the interviewer utilizes this information in order to make sure that the start-up level is set appropriately, not too high or low. The warm-up may also offer some indication as to which topics are familiar and/or interesting to the interviewee.

In the interview phase, the interviewer takes up discussion topics based on the interview protocol. These "topics" are actually areas of discussion, to allow for the dynamic nature of the interactional situation. There are usually 3-4 structured topics on a protocol. The use of the protocol assists in setting up the appropriate task levels for the

elicitation of a representative sample at the skill levels in question. Further, it aids in establishing consistency and, therefore, reliability across results. Of the topics on the protocol, one may happen to be the interviewee's work or area of special interest but other topic areas must also be included. Questions, comments and other prompts from the interviewer should encourage the interviewee to discuss a topic extensively and in depth, e.g. to abstract, support opinions, present hypotheses and counter-hypotheses, offer alternative explanations, clarify, etc. In approximately 10-12 minutes, the interviewer should have established the skill level of the interviewee and start to wind-down the interview with one or two easier questions or interaction initiations.

In the paired discussion phase, the interviewer invites interviewee A back into the room and starts off by asking which topic they have selected for their discussion. The interviewer then observes the discussion without taking part in it, unless such should become necessary, e.g. because of domination of the discussion by one member of the pair. The interviewer may need to stress non-involvement in the discussion (after the interview phase) by explicitly stating that the task is carried out entirely between A and B. Toward the end of the discussion, about 5 minutes, the interviewer may need to put a question or two to the more silent partner before winding-down and closing the session.

Immediately following this, the interviewer should record comments and marks (scores based on the Assessment Criteria for the Advanced Level—Oral Test) on the form provided by the test center. The appropriate information should be filled in and the form should be signed.

ROLE OF THE INTERVIEWER

The interview phase of LUSI consists of two interrelated aspects, elicitation and assessment. This is in reflection of the dual role of the interviewer, who determines in interaction with the interviewee the final form of the task itself and simultaneously conducts an interview which yields a reliable sample of the interviewee's skill level. It goes without saying that this sample provides the basis for evaluation. On one side of the coin, this entails a certain volume of speech elicited. On another, especially in reference to the advanced level, this means "quality", i.e. the topics must have adequate variation and the functions enough challenge to reveal patterns of strengths and weaknesses. If the interviewee is simply allowed to talk about "everyday subjects of conversation" or "my own interests", then oral resources outside these narrow areas remain unexplored. And then, it is impossible with any degree of confidence to assign a skill level. On the other hand, an entire interview beyond the interviewee's, e.g., lexical resources or on uninteresting topics yields an impression of skill which is unfavorable.

The level of difficulty of the questions/interaction initiations put forward by the interviewer depends on the impression, gained through dynamic feedback, of the language abilities of the interviewee. For this reason the interviewer must be very familiar with the Assessment Criteria for the Advanced Level—Oral Test. Nevertheless, even with this knowledge in hand, the interviewer has to beware of relying on a first impression of overall fluency in determining the entry level of difficulty. The lexical range and the interaction skills must always count in the assessment.

Further, in considerations of communicative competence, interaction always entails negotiating intended meanings, i.e. adjusting one's speech to the effect one intends to have on the listener. Therefore, it includes the response of the other and sensitivity to possible misunderstandings. Clarifying one's own and the other's intentions and arriving at the closest possible match between intended, perceived, and

anticipated meanings is part and parcel of this dynamic. This operates in both directions, interviewer to interviewee and vice versa.

Moving through levels of difficulty, skill levels, should be carried out by the interviewer in a way which does not overwhelm the interviewee. In moving between levels the interviewer attempts to determine the basic skill level of the interviewee. This is accomplished by carefully controlling the lexical, prosodic, interactional, discursive, etc. output of the interviewer. When operating within a particular skill level, attempt should be made to parallel, i.e. move to subtopics within this framework which allow for the determination of the range of coverage of the interviewee's skill within that particular skill level. There must be enough of this "movement" in order to determine whether isolated instances of merits or deficiencies in grammar, vocabulary, etc. are simply that, or whether they are, after all, representative of the skill level of the interviewee.

Parallel movement on a skill level basically means that the interviewer attempts to ask questions or make initiations which deal with material which could be construed as similar to the topic under discussion. The idea here is that these questions/initiations are not carried out with the intent of discovering a higher level of proficiency from the interviewee. Parallel movements are intended to ascertain how solidly the interviewee fits within that skill category, whether the interviewee simply happened to make a good statement/interact in a complex manner or whether this type of behavior is representative of a broad range of the interviewee's skills. Such interviewer strategies as moving to subtopics within a framework or probing for a similar type of idiomatic usage (corresponding to the usage just elicited) are representative of this parallel movement.

The rating scale, i.e. the Assessment Criteria for the Advanced Level—Oral Test, is based on hierarchy. This

scale assumes that language use is best assessed holistically. Nevertheless, specific factors contributing to linguistic performance are outlined within it. Furthermore, the scale presumes that facility with a language increases exponentially. This is an expression of the concept that there are greater leaps of performance involved between higher levels than that assumed at lower levels. The rating scale assumes that a performance at any level subsumes the criteria of the levels below it. And it should be pointed out that only sustained performance at a particular skill level is sufficient for rating at that level. The borders between levels always constitute problematic areas for the interviewer. Therefore, attention should be given to them.

The phase of determining the skill level follows seamlessly in a dynamic, iterative process, with the interviewer attempting to elicit higher levels of performance. Once the basic skill level is established, the interviewer should move into linguistic probes which attempt to discover the upper limits of the interviewee's proficiency, i.e. the patterns of weakness. The interviewer continues probing for consistent handling of challenges. The interviewee neither fully succeeds nor completely fails to meet the conversational challenges presented by the probes. However, the level the interviewee can sustain is the level of proficiency.

Nevertheless, this process should not cause consternation in the interviewee. The interviewer must remember to maintain a friendly, positive atmosphere. Further, the interviewer must not be affected by the factual content or accuracy of the statements made by the interviewee.

Please see Appendices 4 and 5 for further practical material on the area of the work of the interviewer. These Appendices give specific guidelines and procedures for the interview situation.

In other words, the interviewer's manner, mode of behavior, facilitates the elicitation of the candidates best performance and the awarding of an accurate rating. This is accomplished by balancing friendliness with neutrality and the interview format with natural conversation. When the interview elicits a clearly ratable sample, the interviewer has succeeded.

APPENDIX 4

NATIONAL CERTIFICATES: INTERVIEWER'S ASSESSMENT GUIDE

NATIONAL CERTIFICATES

ADVANCED LEVEL: SPEAKING (THE FACE-TO-FACE SUBTEST)

INTERVIEWER'S ASSESSMENT GUIDE

GENERAL

This assessment guide aims to help interviewers to rate test takers' speaking skills and to ensure that assessment is comparable between the languages tested in the National Certificates system. The guide includes:

- advice on how the interviewers should assess test takers in the face-to-face subtest, and on writing comments for the assessors
- comments on the assessment criteria
- comments on the proficiency levels (5 - 8) and on the preliminary benchmarks for speaking

The subtest of speaking at the advanced level consists of two parts: 1) speaking tasks that take place in a language laboratory, and 2) face-to-face part which, depending on the language tested, can comprise various types of tasks such as interview / discussion, pair work, or presentation. (Note: On the basic and intermediate levels, speaking is tested only in the language laboratory, except in the Finnish language where either a face-to-face or tape-mediated version of the subtest is used.)

Some important terms:

<u>(overall) rating for speaking</u>	=	the combination of analytic ratings for speaking (this will appear on the certificate)
<u>analytic rating</u>	=	rating given for an individual feature of speaking proficiency assessed (i.e. assessment criterion), such as pronunciation
<u>test battery</u>	=	the whole test containing all five subtests
<u>subtest</u>	=	one part of the whole test battery, such as the reading comprehension subtest

INTERVIEWERS

- an interviewer is a person who interviews / discusses with testees in the face-to-face subtest
- the interviewer also
 - writes comments on the test taker's performance for the assessors
 - assesses testees' performance immediately after the test
- interviewers are experienced language teachers who have been trained to carry out the interviewing and assessment; usually interviewers are also native speakers of the language tested (there is variation between languages in this matter)

COMMENTS ON THE TEST TAKER'S PERFORMANCE

The interviewer should write comments on the test taker's performance on the assessment sheet. These are important for the assessors who rate the performance on the videotape. Among the most useful comments are those relating to the candidate's particular strengths and weaknesses. Also, if the candidate is very good (at level 7 or 8), we hope that the interviewer gives reasons for the particular overall rating for speaking, e.g. why the candidate gets 7 instead of 8.

HOW TO GIVE THE OVERALL RATING FOR SPEAKING

Although not absolutely necessary, we hope that all interviewers give their analytic and overall ratings for the candidates they interview.

When rating the test takers, the interviewer should follow these guidelines:

- the interviewer gives one overall rating for speaking to each testee, but the rating is not given directly (holistically), rather it is combined from analytic ratings. An attempt is thus made to ensure that all relevant assessment criteria are considered when giving the overall rating.
- the interviewer rates the testee's performance with the help of the descriptions of performance found in the rating scales:
 - A) first analytically, that is, the interviewer gives a rating separately for each assessment criterion:
 - 1) pronunciation
 - 2) vocabulary & grammar
 - 3) fluency
 - 4) linguistic appropriacy and discourse skills
(but not for 'clarity of performance' unless the face-to-face part contains a presentation task)
 - B) after this, the interviewer combines the analytic ratings into one overall rating for speaking

- the overall rating for speaking = the most common analytic rating (a rough 'mean' of the ratings)
 - if there is doubt, the 'Overall description of oral skills' found in the beginning of the rating scale may help to decide the rating, although it is just a general description, not an assessment criterion
 - if the interviewer wants to give a higher / lower overall rating than that indicated by the rough mean of the analytic ratings, he or she has to justify this decision and write down his/her reasons on the assessment sheet for the second assessor
- This may be necessary, for example, if the test taker's pronunciation is very peculiar. Another example might be a test taker who somehow fails to fulfill the requirements for the advanced level in a way that may not have been specified in the assessment scales (e.g. he/she is not able to discuss at the necessary level of abstraction or elaborate / justify opinions).

The interviewer should consider the test taker's performance in all parts of the face-to-face subtest when giving the analytic ratings and the overall rating. Thus, the interviewer need not give separate ratings for each part of the subtest (if it has more than one part), but the ratings should be based on the test taker's average performance across all tasks. The test taker's performance can vary, e.g. in the interview he may be very fluent, but hesitate and struggle a lot in the paired discussion. The interviewer must decide the test taker's average performance level (e.g. in fluency, in this example). The following principle may be of help:

- If it is evident that the test taker did poorly in one part of the subtest because of something that was not his/her fault (so that the test taker could not demonstrate his/her ability to speak), the ratings should be based on those parts of the subtest where there were no such problems. For example, it may happen that one of the participants in a pair work dominates discussion too much and leaves the other test taker little room to talk. In this case, the interviewer should ignore the (poor) test taker's performance in the pair work section and rate the person only on the basis of the other tasks. This should, however, be reported in writing to the (second) assessor.

OTHER THINGS TO CONSIDER WHEN RATING THE TEST TAKERS

Before the face-to-face subtest the interviewer should check that the video and audio equipment are in place and that they work, and also locate a person in the test centre to whom to turn to in case of technical problems (the test centre is responsible for providing the recording equipment and tapes, and seeing to it that they work, as well as having an assistant available; however, the interviewer should also personally check the equipment, just to be sure).

During and after the test:

- the interviewer should not take notes nor mark ratings on the assessment sheet while discussing with the test taker because this will interfere with the discussion and may seem unnecessarily threatening to the test taker
- during tasks that do not require the active participation of the interviewer, he/she may of course take notes etc. (e.g. during the paired discussion)
- the interviewer gives the analytic ratings and the overall rating for speaking immediately after the test taker(s) is/are out of the room, or immediately after the last test taker has left the room (depending on the timetable or other practical considerations), while the interviewer can still remember the test taker's performance clearly
- the interviewer can view the video recordings to check his/her ratings, if necessary
- after filling in the assessment sheets for all test takers, the interviewer leaves the sheets and the video tapes at the testing centre to be forwarded to the Language Centre for Finnish Universities and further to the assessors

COMMENTS ON THE ASSESSMENT CRITERIA

Next, each assessment criterion will be examined in turn, and some matters relating to the descriptions used in the scales and to practical assessment will be presented.

The assessment scale consists of six subscales: one for an overall description of oral skills and five for the assessment criteria.

Overall description of oral skills

- this is not a real assessment criterion but a general description of speaking skills at levels 5 - 8
- this can be used by the interviewer and assessor when giving the overall ratings for speaking, i.e., when they are combining the analytical ratings into the overall rating

Pronunciation

- includes both the individual sounds and the prosodic features (intonation, stress, rhythm)
- if the test taker masters one aspect of pronunciation clearly better than another (e.g. individual sounds vs. prosodic features), the rating for pronunciation is a 'mean' of the mastery on these aspects (the interviewer or assessor may decide to weigh one more than the other, but the reason for this should be explained in writing on the assessment sheet)
- the test taker may have an accent even at level 8: the most important thing to consider is clarity, and also the mastery of the prosodic features at level 8

Accuracy of structures and vocabulary, and idiomacy

- accuracy and adequacy of vocabulary for the topics and functions at the advanced level: it is important to assess if there is any merit in the test taker's vocabulary (idioms, ability to distinguish nuances, cultural references)
- structures: accuracy, the ability to use complex and more rare structures
- at levels 6 and 7 there should be only occasional comprehension problems due to errors in vocabulary or structures; at level 8 they should be almost completely absent

Fluency

- definition: hesitation, unnaturally long pausing or groping for expressions that is due to lack of language proficiency (hesitation etc. due to conceptual difficulty is quite another matter and should not be penalised when rating fluency)
- note that the above is a narrower definition of fluency than used in many other contexts
- if the interviewer is uncertain of the cause of non-fluency, he/she should attempt to check out if it is the language that is causing the problems (e.g. by asking more about the same topic)
- the use of good, quick circumlocutions demonstrates good fluency, not a lack of it
- even on level 8, an examinee may search for a word once or twice, but this must happen without a clear interruption in the communication

Linguistic appropriacy and discourse skills

- this is often important in successful communication: mere accuracy and fluency may not always suffice
- definitions: **appropriacy**: level of formality, politeness, directness
 discourse skills: turn-taking, topic switching, reactions; starting, maintaining and concluding a conversation
- this criterion is used in the face-to-face subtest and in the language lab (in the mini-situations and the presentation); the tape-mediated tasks carry more weight in the rating because the tape-mediated tasks allow for a better sampling of different situations

Clarity of performance

- this criterion is used only when rating the presentation task where the test taker has to speak at length (monologue) and to organise the items to be presented into a logical whole
- definition / content: organisation, clarity, efficiency and versatility of cohesion markers

COMMENTS ON THE PROFICIENCY LEVELS

One of the most important principles underlying the National Certificate system is the comparability of the tests and ratings in different languages. The proficiency scales, and the more detailed assessment scales that are derived from them, are central to this aim and should be exactly the same in all languages of the system. The assessors in different languages should all agree on the requirements for each proficiency level, otherwise the testing system loses its credibility. Undoubtedly, the standardisation of assessment is, at least initially, hampered by the different popularity of the languages among the language learners, and thus by the different number of advanced level learners in Finland. There are quite a few really good language users in the languages that are used and studied by many learners (e.g. English, Swedish, Finnish), and far fewer in those languages that are studied less often (e.g. Spanish, Russian). However, this must not lead to different standards in different languages (too lenient in some, too stringent in others): the same performance should be rated in the same way in every language.

The proficiency descriptions used in the National Certificate tests, as well as the preliminary benchmark video samples presented during the first training day, are all based on the work and co-operation of numerous participants. The descriptions and benchmarks may not fit in with some interviewers' or assessors' conceptions of what is meant e.g. by "excellent" or "advanced" language proficiency; some might think that they are too demanding, while others might consider them too lenient. Descriptions of proficiency and the examples that should represent them are always a matter of debate and involve compromises between different views. It is crucial, however, that all interviewers and assessors agree to use common guidelines and standards, otherwise the credibility of the Certificate system cannot be maintained.

The following comments hopefully help the assessors and interviewers in assigning the overall ratings for speaking. The comments also include references to the preliminary benchmark examples that were presented on the first training day. These examples are not always entirely appropriate for the advanced level because the topics discussed are often relatively simple and concrete. However, they illustrate the quality of language expected on the advanced level test: if the test taker can speak about the more demanding topics used in the advanced level tests as well as these benchmark test takers, then the test taker should be given the rating that was given to the benchmark person who best corresponds to the test taker.

Level 5

Speakers at this level have clear flaws in their proficiency, especially when they have to discuss more demanding and unfamiliar topics. Their speech is rather clear and intelligible, but there may be several errors or the speech can be non-fluent at times.

The preliminary benchmarks contain two examples of this level. The woman who speaks Finnish represents a low 5. She was rather clear but not very fluent. Her language was quite accurate grammatically, but her pronunciation was clearly foreign at times. The woman speaking English was an example of a good 5. Some of her skills may have been better than 5 (e.g. pronunciation of sounds), but the problems with fluency and inaccuracies drop the overall rating to 5.

Level 6

Speakers who are at this level have a good command of language which is sufficient for the advanced level tasks. Often the proficiency does not however include any clear merits. On the other hand, many candidates at this level have an uneven skill profile, i.e. some aspects of their proficiency are clearly better than other aspects.

The preliminary benchmarks contained a level 6 woman (speaking English) with an uneven skill profile. Her pronunciation (including prosody) was good, as was fluency at times, but when discussing some topics she became clearly non-fluent and began to make mistakes.

The other example, a rapidly speaking man may be at level 6 or 7, depending on the criteria: fluency was very good, but many assessors find his pronunciation unclear, which may be a reason to give him a 6 as an overall rating (he also has pronunciation errors which often go undetected because of his unclear way of speaking). This is a difficult performance to rate.

Level 7

The test takers' proficiency should be very good also at this level, but there can be more flaws than at level 8.

The preliminary benchmarks contained two examples of level 7 speakers. The woman who spoke Finnish and discussed e.g. children's seats in cars had some grammatical errors and 1 - 2 vocabulary errors. This number of errors is, however, acceptable at level 7. Pronunciation was mostly very natural. Fluency was excellent, and vocabulary was apparently quite sufficient for the (non-abstract) topics at hand; she was also able to use colloquial language.

The other example was a woman speaking English who owned a shop. Her only clear flaw was a very Finnish pronunciation of English, which was however clearly intelligible. Her fluency and the command of vocabulary was very good and compensated for the pronunciation, thus the overall rating of 7.

Level 8

This level represents a very high level of language proficiency, and careful consideration is required whenever the interviewer or assessor wants to give this as an overall rating for speaking, because we want to avoid giving the highest rating too easily. Defining the differences between level 7 and 8 will be one of the most important matters for the whole testing system in the near future.

On the other hand, it is important to remember that level 8 does not stand for perfect and flawless language proficiency: occasional inaccuracies and deficiencies are acceptable. The test takers can include those whose proficiency is at level 8 or even at 9; then it might be dangerous to compare them and give the highest rating only to the best candidate even if he/she exceeds the requirements for level 8. The assessor need only consider if the test taker fulfils the minimum requirements for level 8. For example, the young man who spoke English on the preliminary benchmark video tape was clearly at level 8: his language / communication style may have contained some minor flaws, but

especially his vocabulary far exceeded what can typically be expected at this level. A slightly less proficient speaker would have been at level 8, too.

It is likely that there will be a different number of level 8 speakers in different languages in Finland, depending on the number of those who study these languages: accordingly, the highest overall grades for speaking (and for other skills) will be very rare in the less studied languages, but clearly more common in the more popular languages.

When the interviewer or assessor gives the overall rating for speaking, he/she must take into account all the assessment criteria. The overall rating is a kind of rough mean of the analytical ratings. If the interviewer/assessor wants to deviate from this rule and to weight one or two of the criteria more than the others, he/she should write the reasons for this decision.

FEEDBACK IS REQUESTED

The National Certificate tests have been developed in co-operation with a number of experts. Now that the tests are administered for the first time, a large number of new, knowledgeable people joins the testing group, i.e. interviewers and assessors who have experience on teaching and testing languages. We hope that this co-operation will be as fruitful as the design of the first test batteries has been.

Feedback from assessors and interviewers will be invaluable for the further development of the assessment criteria and the rating system in general. We are especially interested in getting your feedback on use of 'linguistic appropriacy' and 'clarity of presentation' for assessment because there is very little research on these criteria in the literature. It has also been rather difficult to design the rating scales with descriptions for these criteria since there is less agreement on what should be included in them than there is for most other criteria.

Also, feedback is needed to decide whether it is necessary to make any language specific changes in the assessment scales. The starting point should be that the scales are the same for all languages but if there might be good reasons e.g. to illustrate the scales with good examples, or to highlight some matters in a different way in the descriptions. Any changes to descriptions will be made later when we have sufficient evidence for their necessity.